Durham E-Theses

Bayes Linear Variance Learning for Mixed Linear Temporal Models

RANDELL, DAVID

How to cite:

RANDELL, DAVID (2012) *Bayes Linear Variance Learning for Mixed Linear Temporal Models*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/3646/

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, Durham University, University Office, Old Elvet, Durham DH1 3HP
e-mail: e-theses.admin@dur.ac.uk Tel: +44 0191 334 6107
http://etheses.dur.ac.uk

Bayes Linear Variance Learning for Mixed Linear Temporal Models

David Randell

Submitted for the degree of Doctor of Philosophy
Dec 2011

Abstract

Modelling of complex corroding industrial systems is critical to effective inspection and maintenance for assurance of system integrity. Wall thickness and corrosion rate are modelled for multiple dependent corroding components, given observations of minimum wall thickness per component. At each inspection, partial observations of the system are considered. A Bayes Linear approach is adopted simplifying parameter estimation and avoiding often unrealistic distributional assumptions. Key system variances are modelled, making exchangeability assumptions to facilitate analysis for sparse inspection time-series. A utility based criterion is used to assess quality of inspection design and aid decision making. The model is applied to inspection data from pipework networks on a full-scale offshore platform.

Bayes Linear Variance Learning for Mixed Linear Temporal Models

David Randell

A Thesis presented for the degree of
Doctor of Philosophy

Statistics Group
Department of Mathematical Sciences
University of Durham
England

December 2011

Contents

List of Figures

List of Tables

Declaration

The work in this thesis is based on research carried out at the Statistics Group, the Department of Mathematical Sciences, University of Durham, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it all my own work unless referenced to the contrary in the text.

Copyright © 2011 by David Randell.

"The copyright of this thesis rests with the author. No quotations from it should be published without the author's prior written consent and information derived from it should be acknowledged".

Acknowledgements

I would like to thank the Engineering and Physical Sciences Research Council and Shell Global Solutions for joint sponsorship of this work.

I would also like to thank my supervisor Michael Goldstein of University of Durham, Philip Jonathan of Shell and all my friends and family who have pushed me to finish this thesis.

Chapter 1

Complex industrial systems

1.1 Complex industrial systems

Industrial processes contain large complex structures made up of many interconnected parts or components. Such systems need careful monitoring to assess performance. Deterioration of these systems, due to processes such as corrosion, can lead to reduced performance, and cause system failures and breakdowns. These system failures can be extremely financially and environmentally costly and are a potential risk to safety. Inspection is usually carried out to assess the current state of repair of the system and where necessary perform maintenance.

The aim of this thesis is to consider aspects of a tractable approach to modelling and learning about complete systems. We present methodology to update model parameters including uncertainties for irregular, incomplete, short time series. Then we show how to design inspection schemes to improve efficiency and effectiveness of inspection, minimising potential losses whilst designing to learn about means and variances.

1.2 Inspection methods

In many cases industrial process are extremely slow and costly to shut down. As a result, inspections are expensive and not easy to carry out. Non-destructive testing (NDT) or non-invasive inspection (NII) methods are normally used to inspect these

systems, i.e. components are inspected in-situ without damaging or affecting the operation of the system. Due to the materials involved, corrosion and fatigue are often internal and cannot be seen from outside. Most NDT inspection techniques involve sending an energy source (e.g. sound or electromagnetic) into a material, and using a detector to measure that energy. Differences or flaws in the structural integrity of the material will then affect the energy source as it moves through the material, giving a picture of its internal structure. ASNT [2009] gives an introduction to NDT.

Ultrasonic testing involves sending sound waves into materials to check for internal flaws. The time the sound takes to be sent out and reflect back to the detector gives an idea of component wall thickness. It can be used to detect very small flaws in the object and doesn't affect the operation of the component. Krautkrämer and Krautkrämer [1990] provides a very detailed explanation of the process of ultrasonic testing. Time of Flight Diffraction (TOFD) is a fast method of implementing ultrasonic inspection, which uses a pair of probes sending signals down the length of a component, allowing quick inspection over a large area.

Radiographic testing involves using gamma rays to penetrate materials to give an image of the internal structure of the material. The density of the material affects the rate of diffusion of the radiologic source and thus affect how clearly various features and flaws show up and therefore requires calibration.

1.3 Corrosion

Corrosion is a chemical process where a metal reacts with chemicals around it producing a new and less desirable material. The most commonly known form of corrosion is rusting, which consists of iron or steel (or other iron alloys) reacting with oxygen to form ferrous oxides. Iron in the presence of water and oxygen will form Iron Hydroxide,

$$2Fe + O_2 + 2H_2O \rightarrow 2Fe(OH)_2$$

this Iron Hydroxide will then react further with oxygen to form hydrated iron oxide (rust), [Nimmo and Hinds, 2003].

$$4Fe(OH)_2 + O_2 \rightarrow 2Fe_2O_3.H_2O$$

Rust is porous, and allows air though, enabling corrosion to continue further inside the iron beyond the rust. Aluminium is a metal which naturally forms a protective barrier that insulates it from corrosion. Aluminium is a very reactive metal which in the presence of air quickly oxidises, just like iron. However, unlike iron, this oxide is not very reactive and insulates the inside of the material. This process, known as passivity, can be replicated in other materials such as iron/steel, giving the metal a protective coating that is non-reactive, protecting the metal from corrosion. Stainless steel is one such example which is an alloy of iron that includes chromium to help protect it from corrosion. Talbot and Talbot [1998] gives an overview of corrosion and protection strategies. There are various different types of corrosion.

- *uniform corrosion:* the gradual degradation of metals, such as rusting, which is normally fairly consistent and predictable.
- *pitting corrosion:* a localised type of corrosion, where the rate of corrosion in particular places is significantly higher than other areas. This type of corrosion typically occurs in places where protective coatings have broken and can lead to cracks and failure of components.
- *galvanic corrosion:* where 2 different metals are in contact with each other and have different reactivity. One metal can give up electrons and ions to the other metal.
- *stray current:* an electric current acting on a metal can turn it into a electrode and cause it to leech material away.
- *microbial attack:* certain fungi and microbes can cause damage and corrosion to metals.

An illustration of a pipework section which shows evidence of uniform and pitting corrosion is shown in figure 1.3. The temperature of metals also affects how fast

Figure 1.1: Internal corrosion of a typical pipework section

things corrode. In a high temperature environment, metals become more reactive and thus more susceptible to corrosion. Uhlig [1971] provides a detailed summary of types of corrosion damage that can occur.

1.4 Offshore platform data

Throughout this thesis we will illustrate ideas and methodology with an example of a complex industrial system: an offshore oil platform, figure 1.4. On an offshore platform, there are a large number of complex pipework networks. These can be broken down into many corrosion circuits. Within an individual corrosion circuit, there will be many interconnected components. A component, c, may be considered as a small region/section of a corrosion circuit which can be replaced and behaves in a consistent way with respect to its corrosion properties. A schematic of one such corrosion circuit is shown in figure 1.3 with pipework structure is shown by lines and individual components marked and numbered. The schematic shows a pipework system where components occur at irregular intervals. There are several branches

Figure 1.2: A typical offshore platform

and sub-branches along the length of the corrosion circuit. It seems reasonable to think that components which are corroding in the same system, subject to similar conditions will behave in similar ways so we want to be able to include this feature whilst modelling. The strength of relationships between components will be related to how close together they are. Although components may be physically closer to one another the important factor controlling relationships between components is the notion of distance along the pipe. We will therefore consider a distance metric as the number of components between components along the pipe. Multiple Corrosion circuits, although not necessarily linked, also perform similar tasks and also may be related.

1.4.1 Inspection of an offshore platform

Inspection and maintenance on offshore platforms is a difficult task. As a result it is very rare that complete systems will be inspected. Inspections are carried out at relatively large time intervals, t, (months or years apart) and not carried out on a regular basis. So we end up with short irregular time series for individual components within the system. However we do have large numbers of components

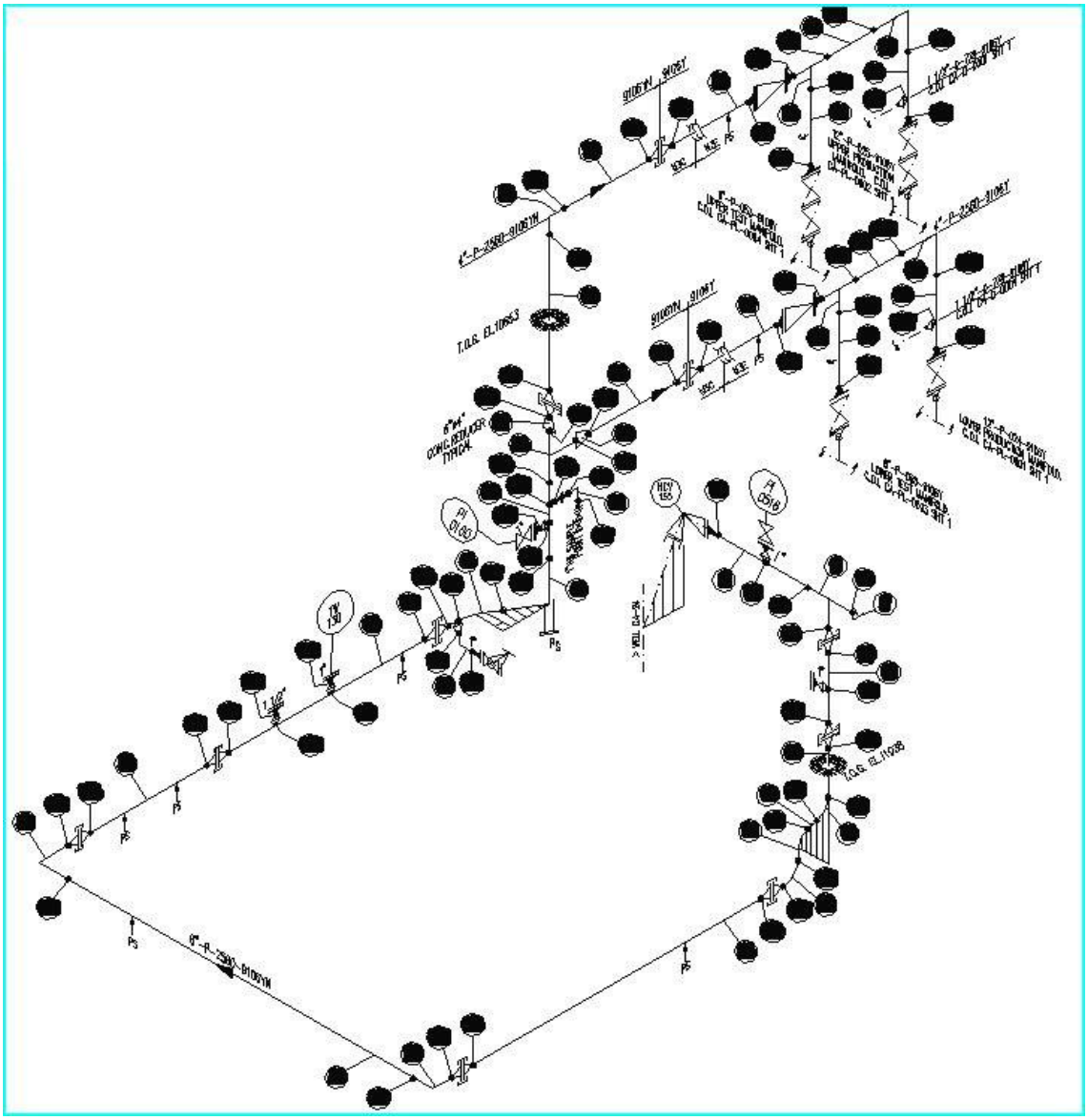

Figure 1.3: A typical schematic diagram of a corrosion circuit. Dots on schematic represent individual components

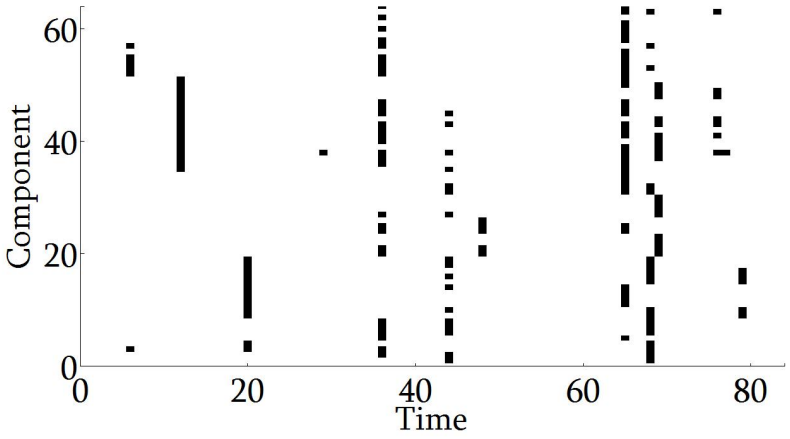

Figure 1.4: A typical inspection design for the offshore application, consisting of 64 components over 83 time points. Black lines correspond to 174 observations of the system.

so sharing information across components is important. When components are inspected, generally inspectors are most concerned with where the component is at its most vulnerable. Consequently only one measurement is usually recorded for a whole component, the minimum wall thickness, since this is where the component is most likely to break.

For the current application, we will model a system of 4 corrosion circuits consisting of a total of 64 pipe-work weld components. Historical data are available for component minimum wall thickness, obtained during inspection campaigns using non-intrusive ultrasonic measurements for the period 1998 - 2005. The data is given in a table in the appendix A. Based on the frequency of observations and the requirements for inspection planning, we select a monthly time-increment for modelling; the historical period therefore consists of 83 time points. The actual historical inspection design is given in Figure 1.4. From the figure it is clear that inspections are typically incomplete and irregularly spaced in time. A total of 174 observations of the system are available.

1.4.2 Synthetic data

Applying methodology to real-life examples is important. However it can be difficult to analyse performance since we don't know what the "real" answer should be. The use of diagnostic tests gives us some idea of whether things are behaving as expected. To illustrate ideas then we will also be using simulated data sets similar in set up to the real data, where we know the underlying parameters. We can then quantify exactly how well our methodology is performing. Discussion of the use of simulators for large systems is given in Craig et al. [2001] and Goldstein and Rougier [2006].

1.5 Contribution of the thesis

The original contribution of the thesis is in presenting a methodology for tractable modelling of large systems of short time series using incomplete and irregularly spaced correlated time series. The use of second order exchangeability assumptions provides a structure to correlations and allows interference on both mean and variance quantities.

1.6 Thesis outline

In chapter 2 we introduce some of the key theoretical concepts used in the thesis and examine other current approaches in this area. We then discuss a modelling approach for complex systems in chapter 3, modelling the systems as a combinations of a global and local system. In chapter 4 we look at how to use historical data to update our understanding of the system using Bayes Linear analysis. We also look at the specification of prior beliefs and also diagnostics to check consistence and coherence of data. We then move on to look at the problem of learning about variability. In chapter 5 we look at Bayes linear variance learning for a linear growth DLM, using exchangeability assumptions. In chapter 6 we then generalise to update the global effect variances for the model described in chapter 3; learning about variances in the case of a non linear observation equation. We look at using the Mahalanobis Distance for parameter estimation including variance learning in chapter 7. A candidate set of

parameter values are chosen, each of which is compared with observations using the Mahalanobis distance. We then consider how to design efficient inspection schemes in chapter 8. A utility based criterion is used to assess designs. Designs which inspect components about which there is a high probability of failure or about which there is high uncertainty are favoured. Discussion of the methodology and some areas of further study is given in chapter 9. In appendix C is a table of with an explanation of the notation used in this thesis.

Chapter 2

Theoretical background

We now introduce some of the key concepts we will use in this thesis as well as reviewing current approaches in the areas of corrosion modelling, inspection and Bayesian approaches dealing with large problems. We begin by examining existing literature before presenting an introduction to Bayes linear statistics and second order exchangeability.

2.1 Literature review

We begin by reviewing existing literature on the subjects of corrosion, modelling, inspection and Bayesian estimation. We are considering large industrial systems made up of many interconnected related parts or components. We will model complete systems and update beliefs about the system state including uncertainties in situations where we have irregular, incomplete, short time series.

2.1.1 Corrosion modelling

Corrosion science is an area in its own right with a vast body of literature. Nimmo and Hinds [2003] present a beginners guide to the chemical process of corrosion, whilst Talbot and Talbot [1998] and Uhlig [1971] provide a more detailed explanation of corrosion science.

Models are an attempt to quantify the physical word mathematically. Modelling corrosion helps to improve understanding of the system and aid prediction. Industry

guidelines such as Health and Executive [2002] and ASTM-Standard-G16 [2004] talk in very general terms about the statistics of modelling corrosion. Zhang and Mahadevan [2000] outline mathematical expressions for initiation and evolution of different corrosion mechanisms, including pitting and cracking. Estes and Frangopol [2001] present a cost based approach for maintenance planning, for deterioration of structures with application to concrete bridges.

Some approaches to corrosion modelling consider corrosion as an extreme value problem. Coles [2001] provides a good introduction to the statistical modelling of extremes and Kotz and Nadarajah [2000] gives an overview of extreme value distributions. Generalised Extreme Value distributions are used to model corrosion in Glegola [2007]. Scarf and Laycock [1996] give a review of extreme value models used in corrosion modelling. Caleyo et al. [2009] use Monte Carlo simulation to investigate external corrosion pit depth, and pit growth rate, in underground pipelines. Qin and Cui [2003] provide a corrosion model describing the corrosion process on steel structures. Another examination of corrosion in pipelines is given in Hawn [1977], where extreme value analysis is used in the prediction of maximum pit depth.

2.1.2 Inspection and maintenance

As well as modelling corrosion, inspection and maintenance planning is an important area of study. Kallen and van Noortwijk [2005] look at maintenance planning, considering situations where inspections are not perfect. A number of authors discuss the inclusion of inspection data and expert judgement, within a risk-based inspection framework. For example, Faber and Sorensen [2000] present an approach to estimating the condition of systems, for inspection planning purposes, using a combination of inspection observations and expert judgement. Straub [2004] describes generic approaches to risk-based inspection of steel structures. Incomplete data is a feature of inspection of complex industrial systems. Gasemyr and Natvig [2001] present a Bayesian approach to modelling in situations of partial inspections. It is not always known whether corrosion has been initiated within a system. Kuniewski et al. [2009] present a method for compliance sampling, deciding whether or not corrosion has initiated and how much of a system needs inspecting.

The majority of these approaches to inspection look at individual system components. Alternatively, when modelling larger systems, they consider cases where components are independent, effectively confining the problem to a univariate case.

Our approach models the entire system and consequently, to retain tractability, a less complex method of modelling is needed. Little et al. [2004a] use a multivariate dynamic linear model to characterise the corrosion of large industrial storage tanks, using observations of component minima, and suggest approaches to optimal inspection planning. Little et al. [2004b] describe the application of a spatio-temporal dynamic linear model to modelling the corrosion of an industrial furnace, using Bayes Linear updating. Empirical distance-based estimates for covariances of dynamic linear model observations and system variances are used, and optimal inspection planning based on heuristic criteria is considered. Shaddick and Wakefield [2002] use a hierarchical dynamic linear model, modelling daily multivariate pollutant data, for multiple sites. Alternatively, G.Hardman [2007] presents a utility based approach to evaluate costs of inspection designs, to give a monetary criterion.

2.2 Bayes linear analysis

Bayesian theory allows incorporation of prior beliefs and expert judgement, together with observational data. It can deal with very complex types of modelling. However quantifying prior beliefs can be difficult especially when dealing with experts without a background in statistics. O'Hagan et al. [2007] deal with the area of elicitation and describe approaches to help extract expert judgements. Farrow [2003] examines elicitation and building of subjective covariance structures, for large systems.

2.2.1 The Bayes linear approach

Modelling large and complex systems by full Bayesian analysis can be difficult. Even in small problems, with few sources of uncertainty, it can be difficult to assess a satisfactory full joint prior probability specification over all of the possible outcomes. Unrealistic, simplifying prior assumptions are often made (e.g. conjugacy) to make inference possible [Raiffa and Schlaifer, 1961].

Improvements in modern computing has meant that stochastic simulation approaches such as Markov Chain Monte Carlo Methods have become more prevalent Robert and Casella [2004]. The advantage of these methods is that given enough computational time, full and complex posterior inferences can be made. However, in large problems such as complex industrial systems, there may be hundreds of relevant sources of uncertainty about which prior judgements are made. In such problems, it is arguably impossible for us to carry out a full Bayesian analysis. If such a full prior specification were possible, it would often be the case that the specification was too time consuming and too difficult to check. Further, the resulting Bayesian analysis can often be extremely computer intensive. The Bayes linear approach is particularly appropriate whenever the full Bayesian approach requires an unnecessarily exhaustive description and analysis of prior uncertainty.

The Bayes linear approach can be viewed as either (a) offering a simple approximation to a full Bayesian analysis, for problems where the full analysis would be too difficult or time consuming, or (b) complementary to the full Bayes analysis, offering a variety of new interpretative and diagnostic tools which may be of value whatever our viewpoint, or (c) a generalisation of the full Bayesian approach where the artificial constraint that requires a full probabilistic prior specification is lifted. The Bayes linear approach is particularly helpful for design problems where we need to search over a large design space. In these cases other approaches would become computationally intractable.

Prior beliefs are specified in terms of means, variances and covariances. A brief summary of the Bayes linear approach is given Goldstein [1998] with a detailed treatment in Goldstein and Wooff [2007]. Farrow and Goldstein [1993] discuss Bayes linear methods, for grouped multivariate repeated measurement studies, with application to cross-over trials. Modelling large systems can be computationally intensive; Bayes linear analysis can be used effectively in large simulator problems. Craig et al. [2001] discuss the use of large scale simulators and use a Bayes linear approach to forecasting. Goldstein and Rougier [2006] discuss the calibration of large simulators, based on adjustment of observational data. The Bayes linear approach, provides a computationally efficient method for updating beliefs for problems where

a full Bayesian approach would be too difficult or time consuming. The basis of Bayes linear analysis, comes from the idea that expectation, not probability, should be the primitive quantity. The concept of prevision (or expectation) is explored by de Finetti [1974]. Using the idea of the subjective pricing of gambles, the price you are willing to pay for a gamble is your prevision. This price is then unique to "You", hence, subjective. Whittle [1992] provides a non-subjective approach to probability theory, derived from expectation not probability.

Learning about variances within models is another important area. This is an inherently difficult task, since multiple observations are needed to understand variability. Wilkinson [1997] discusses variance learning for a univariate linear growth dynamic linear model, and Wilkinson and Goldstein [1997] describe Bayes linear covariance matrix adjustment for a multivariate constant dynamic linear model. However, these examples are concerned with long time series and complete data.

Short irregular time series data for large numbers of components is similar to longitudinal studies where modelling covariance is important. Rawwash [2005] presents a method of covariance matrix estimation for use when data are incomplete. Diggle and Verbyla [1998] and Wu and Pouramadi [2003] both consider non-parametric approaches to estimation of large covariance matrices.

2.2.2 Adjusting beliefs

Suppose we have a collection of vector quantities, X, where $X = \{X_1, X_2, \ldots X_N\}$. We then observe a collection, D, where, $D = \{D_0, D_1, D_2, \ldots D_T\}$ (where $D_0 = 1$) and wish to adjust our beliefs about X, given observations, D.

The adjusted expectation of X is defined to be, $E_D(X)$, the linear combination of the D which minimises:

$$E((X - \sum_{i=0}^{k} h_i D_i)^2),$$

over all possible choices of h. This is also known as the Bayes linear rule for X given D. This can be simplified to give:

$$E_D(X) = E(X) + \mathrm{Cov}(X, D)[\mathrm{Var}(D)]^{\dagger}(D - E(D)) \tag{2.2.1}$$

where $\mathrm{Var}(Y)^{\dagger}$ is the inverse of $\mathrm{Var}(Y)$ if invertible; If not, a generalised inverse,

such as the Moore-Penrose pseudo inverse should be used. The adjusted variance, $\mathrm{Var}_D(X)$, is given by:

$$\mathrm{Var}_D(X) = \mathrm{Var}(X) - \mathrm{Cov}(X, D)[\mathrm{Var}(D)]^{\dagger}\mathrm{Cov}(D, X) \tag{2.2.2}$$

and may be considered as the mean squared error of the estimator $E_D(X)$. The resolved variance,$\mathrm{RVar}_D(X)$, is then given by:

$$\mathrm{RVar}_D(X) = \mathrm{Cov}(X, D)[\mathrm{Var}(D)]^{\dagger}\mathrm{Cov}(D, X) \tag{2.2.3}$$

which is the amount of variation of X, which is explained by D.

When considering design calculations, i.e. optimal inspection schemes, we will need the ability to examine large numbers of designs and compute beliefs quickly. Due to the linear nature of the equations, Bayes linear analysis gives the ability to update our beliefs quickly, even for high dimensional problems. We also only need to specify a partial belief structure of first and second order quantities; that is means, variances and covariances. We therefore avoid the need to assume specific distributional forms. In the context of large industrial systems, we also need to be able to specify relationships between components to be able to share information. We do this using exchangeability.

Goldstein and Wooff [2007, page 61-63] discuss the interpretation of adjusted beliefs, which can be viewed as giving simple tractable approximations to a full Bayesian approach. More formally, de Finetti defines expectation as the value λ that minimises the following penalty function (with $c > 0$ denoting unit loss):

$$L = c[X - \lambda]^2$$

In this case adjusted expectation $E_D(X)$ extends the definition of expectation to the case where there are a range of values of λ:

$$L_D = c[X - \sum_{i=0}^{k} \lambda_i D_i)]^2$$

Conditional expectation $E(X|D)$ can be expressed as the penalty function:

$$L_D = cD[X - E(X|D)]^2$$

consequently; adjusted expectation can be expressed as:

$$E_D(X) = \sum_{i=1}^{k} E(X_i|D_i)D_i$$

In the case where D is a partition so that each D_i is 1 or 0 and $\sum_i(D_i) = 1$ then adjusted expectation and conditional expectation are identical.

2.3 Exchangeable events

Consider a situation where we have two dice, die A and die B. We roll both dice and consider the vector of face scores. Suppose that this vector is unaffected by the order in which we roll the dice, i.e. whether we roll dice A first then dice B or the other way around, our beliefs about the vector of scores remain unchanged. This idea is known as exchangeability.

The concept of exchangeable events is a crucial component of the subjective theory of probability. In essence, exchangeability assumptions in a subjective analysis provide a similar mathematical framework to independence assumptions in classical inference [de Finetti, 1974]. Goldstein [1986] discusses exchangeable belief structures, and presents a way of representing exchangeable quantities through mean and residual elements. Goldstein and Wooff [1998] describe a process for the Bayes linear adjustment of exchangeable beliefs. Multivariate partial exchangeability is considered in Shaw [2000], who examines the idea of co-exchangeable systems. In the context of Bayes linear analysis, where only partial beliefs need to be specified, we can restrict exchangeability to first and second order quantities (means, variances and covariances).

2.3.1 Second order exchangeability

Definition 2.3.1 A collection of vector quantities $X = \{X_1, X_2, \dots\}$ is second order exchangeable if our beliefs about the first and second order specification are invariant under permutation of X [Goldstein and Wooff, 2007, page 195]

So again, if we have two dice, we roll both dice and consider the vector of face

scores. Our beliefs regarding the mean and variance of that vector of face scores, are unaffected by the order in which we roll.

Given a collection of vector quantities $X = \{X_1, X_2, \dots\}$ which are second order exchangeable, each has the same mean, variance and covariance between them i.e.:

$$E(X_i) = \mu \quad \text{Var}(X_i) = \Sigma \quad \text{Cov}(X_i, X_j) = \Gamma \quad i \neq j$$

This leads to the following representation theorem for second order exchangeable quantities as given in [Goldstein and Wooff, 2007, page 186] .

Theorem 2.3.2 Given a collection of vector quantities $X = X_1, X_2, \dots$ which are an infinitely exchangeable sequence with:

$$E(X_i) = \mu \quad \text{Var}(X_i) = \Sigma \quad \text{Cov}(X_i, X_j) = \Gamma \quad i \neq j$$

Then we may express each X_i as:

$$X_i = \mathcal{M}(X) + \mathcal{R}_i(X)$$

where $\mathcal{M}(X)$ is a random vector known as the population mean with:

$$E(\mathcal{M}(X)) = \mu \quad \text{Var}(\mathcal{M}(X)) = \Gamma$$

and the collection $\mathcal{R}_i(X)$ are known as the residuals and are also second order exchangeable with:

$$E(\mathcal{R}_i(X)) = 0 \quad \text{Var}(\mathcal{R}_i(X)) = \Sigma - \Gamma$$

Each $\mathcal{R}_i(X)$ and $\mathcal{R}_j(X)$ are mutually uncorrelated for $i \neq j$ and each $\mathcal{R}_i(X)$ is uncorrelated with $\mathcal{M}(X)$.

2.4 Adjusting exchangeable quantities

Using exchangeability assumptions allows us to specify relationships between objects simply through means, variances and covariances. Combining that with the Bayes linear approach, we have a simple way of adjusting our beliefs given observational data. Specifying belief structures through the use of exchangeability greatly simplifies the Bayes linear adjustment.

Assume X is second order exchangeable with:

$$E(X_i) = \mu \quad \text{Var}(X_i) = \Sigma \quad \text{Cov}(X_i, X_j) = \Gamma \quad i \neq j$$

then from the representation theorem 2.3.2 we can express each X_i as

$$X_i = \mathcal{M}(X) + \mathcal{R}_i(X)$$

where:

$$\begin{aligned} E(\mathcal{M}(X)) &= \mu & \text{Var}(\mathcal{M}(X)) &= \Gamma \\ E(\mathcal{R}_i(X)) &= 0 & \text{Var}(\mathcal{R}_i(X)) &= \Sigma - \Gamma \\ \text{Cov}(\mathcal{M}(X), X_i) &= \Gamma & \text{Cov}(\mathcal{R}_i(X), X_i)\Sigma - \Gamma \end{aligned}$$

Given this representation we can write down the adjustment equation. Assume we observe X_i, then by equation 2.2.1 we can find the adjusted expectation of $E_{X_i}(\mathcal{M}(X))$:

$$\begin{aligned} E_{X_i}(\mathcal{M}(X)) &= E(\mathcal{M}(X)) + \text{Cov}(\mathcal{M}(X), X_i)[\text{Var}(X_i)]^{\dagger}(X_i - E(X_i)) \\ &= \mu + \Gamma\Sigma^{\dagger}(X_i - \mu) \end{aligned}$$

This is a weighted average between the observation and prior expectation, with the weighting controlled by the correlation between vectors.

When adjusting exchangeable quantities, we don't need to adjust beliefs using individual observations. The sample mean is a sufficient statistic. Also we only need to adjust beliefs about the population mean $\mathcal{M}(X)$.

Theorem 2.4.1 Let $\{X_1, X_2, \dots\}$ be an infinite second order exchangeable sequence. Then the sample mean vector:

$$\bar{X_n} = \frac{1}{n}\sum_{j=1}^{n} X_j$$

taken from a sample:

$$D_n = (X_1, X_2, \dots, X_n)$$

is Bayes linear sufficient for adjusting beliefs about $\mathcal{M}(X)$. Therefore, to learn about any element of D_n we simply need to adjust our beliefs about the mean $\mathcal{M}(X)$ and

for any values X_i, $i > n$:

$$E_{D_n}(X_i) = E_{D_n}(\mathcal{M}(X)) = E_{\bar{X_n}}(\mathcal{M}(X))$$
$$\mathrm{Var}_{D_n}(X_i) = \mathrm{Var}_{D_n}(\mathcal{M}(X)) = \mathrm{Var}_{\bar{X_n}}(\mathcal{M}(X) + \mathrm{Var}(\mathcal{R}_i(X)))$$

see Goldstein and Wooff [2007, page 208].

The adjusted expectation, $E_{\bar{X_n}}(\mathcal{M}(X))$ is therefore given by:

$$\begin{aligned} E_{\bar{X_n}}(\mathcal{M}(X)) &= E(\mathcal{M}(X)) + \mathrm{Cov}(\mathcal{M}(X), \bar{X_n})[\mathrm{Var}(\bar{X_n})]^\dagger(\bar{X_n} - E(\bar{X_n})) \\ &= \mu + \Gamma(\Gamma + \frac{1}{n}(\Sigma - \Gamma))^\dagger(\bar{X_n} - \mu) \end{aligned}$$

Similarly the adjusted expectation, $E_{\bar{X_n}}(\mathcal{R}_i(X))$ is given by:

$$\begin{aligned} E_{\bar{X_n}}(\mathcal{R}_i(X)) &= E(\mathcal{R}_i(X)) + \mathrm{Cov}(\mathcal{R}_i(X), \bar{X_n})[\mathrm{Var}(\bar{X_n})]^\dagger(\bar{X_n} - E(\bar{X_n})) \\ &= 0 + (\Sigma - \Gamma)(\Gamma + \frac{1}{n}(\Sigma - \Gamma))^\dagger(\bar{X_n} - \mu) \end{aligned}$$

This means when we are adjusting beliefs about exchangeable quantities, it is enough to adjust beliefs about $\mathcal{M}(X)$. We also only need to update using sample means. This approach greatly simplifies belief updating.

The adjusted variance, $\mathrm{Var}_{\bar{X_n}}(\mathcal{M}(X))$ is given by:

$$\begin{aligned} \mathrm{Var}_{\bar{X_n}}(\mathcal{M}(X)) &= \mathrm{Var}(\mathcal{M}(X)) - \mathrm{Cov}(\mathcal{M}(X), \bar{X_n})[\mathrm{Var}(\bar{X_n})]^\dagger \mathrm{Cov}(\bar{X_n}, \mathcal{M}(X)) \\ &= \Gamma - \Gamma(\Gamma + \frac{1}{n}(\Sigma - \Gamma))^\dagger \Gamma \end{aligned}$$

The adjusted variance, $\mathrm{Var}_{\bar{X_n}}(\mathcal{R}_i(X))$ is given by:

$$\begin{aligned} \mathrm{Var}_{\bar{X_n}}(\mathcal{R}_i(X)) &= \mathrm{Var}(\mathcal{R}_i(X)) - \mathrm{Cov}(\mathcal{R}_i(X)\bar{X_n})[\mathrm{Var}(\bar{X_n})]^\dagger \mathrm{Cov}(\bar{X_n}, \mathcal{R}_i(X)) \\ &= \Sigma - \Gamma - (\Sigma - \Gamma)(\Gamma + \frac{1}{n}(\Sigma - \Gamma))^\dagger(\Sigma - \Gamma) \end{aligned}$$

Chapter 3

Modelling complex industrial systems using mixed linear temporal models

We now present a generalised approach to modelling the evolution and state of complex systems which are degrading in time. We model the true state of the system, Z. We model the surface of each component c, as a grid of L_c locations, l, for each component in the system, c, over time t. We separate global aspects which affect the whole component from local aspects. This allows us to distinguish between different model characteristics. Through the global effects model, we capture the most important features and model relationships between components. The local effects model captures behaviour of the surface, in more detail.

3.1 Global effects model

3.1.1 Dynamic linear models

We model the global effects through a dynamic linear model (DLM). Harrison and West [1989] provides a excellent reference on the subject of modelling with DLM's. A dynamic model is a statistical model, which adapts to changes in the model parameters over time. This is appealing from a Bayesian point of view since inspec-

tion data can be incorporated into the model from which we learn about the model parameters.

Definition 3.1.1 A constant DLM is given by the model equations:

$$\begin{aligned} Z_t &= F\Theta_t \\ \Theta_t &= G\Theta_{t-1} + \epsilon_{\Theta t} \end{aligned}$$

where at discrete times t, $Z_t = \{Z_{t1}, Z_{t2}, \ldots, Z_{tC}\}^T$ is a $C \times 1$ system state vector over components, where:

- $\Theta_t = (\Theta_{t1}, \Theta_{t2}, \ldots, \Theta_{tN})^T$ is the $N \times 1$ state vector over parameters at a particular time, t,;
- F is a known $(C \times N)$ dynamic regression matrix
- G is a known $(N \times N)$ system state evolution matrix.
- $\epsilon_{\Theta t}$ is the $N \times 1$ vector of system state evolution residuals. Properties of the residuals are given in section 3.1.3.

In general ϵ is used model residual quantities,. However since there are several different residuals within the models of interest a convention is used; the parameter to which the residual refers is referenced using a subscript e.g. ϵ_θ is the residual belonging to the parameter vector θ.

3.1.2 Global effects model

Given a system with C components, we model as above the $N \times 1$ vector of system state, Θ_t, at time, t, where N is the number of parameters (indexed by n) in the state vector for the model. This parameter set may include multiple parameters per component (for example a wall thickness and corrosion rate parameter per component giving $N = 2C$.)

The global effects model does not depend on location, l but does depend on component c. Using the same notation to section 3.1.1 we have:

$$\Theta_t = G\Theta_{t-1} + \epsilon_{\Theta t} \tag{3.1.1}$$

where G is the, $N \times N$, system evolution matrix and controls how the parameters evolve through time. The system evolves with discrete equally spaced time steps t. $\epsilon_{\Theta t}$ is the vector system evolution residual, at time t:.

The correlation in the residual structure control the relationships between components and parameters in the system. When updating beliefs, the residual structure also controls how information is passed across components.

3.1.3 Exchangeable residual structures

We use exchangeability assumptions to describe beliefs about the system evolution residuals. Exchangeability is discussed in section 2.3. We make the following exchangeability assumptions regarding the residual structure.

Exchangeability of system evolution residuals in time

Firstly we assume that system evolution residuals $\epsilon_{\Theta nt}$ for each parameter n, are second order exchangeable in time with:

$$E(\epsilon_{\Theta nt}) = 0 \quad \text{Var}(\epsilon_{\Theta nt}) = \Sigma_{\Theta n} \quad \text{Cov}(\epsilon_{\Theta nt}, \epsilon_{\Theta nt'}) = 0 \quad t \neq t'$$

where $\Sigma_{\Theta n}$ is a scalar variance which is constant in time for each parameter, n. So by theorem 2.3.2 we get a representation, for every parameter $n = 1, 2, ...N$:

$$\epsilon_{\Theta nt} = \mathcal{M}(\epsilon_{\Theta n}) + \mathcal{R}_{nt}(\epsilon_{\Theta n}) \tag{3.1.2}$$

and in this case $\mathcal{M}(\epsilon_{\Theta n}) = 0$ so this reduces to:

$$\epsilon_{\Theta nt} = \mathcal{R}_{nt}(\epsilon_{\Theta n})$$

and we get that:

$$\text{Var}(\mathcal{R}_{nt}(\epsilon_{\Theta n})) = \Sigma_{\Theta n}$$

Relationship of system evolution residuals across parameters

We also assume that for fixed time t, the residuals $\epsilon_{\Theta nt}$ are related across parameters such that:

$$\text{Cov}[\epsilon_{\Theta nt}, \epsilon_{\Theta n't'}] = \begin{cases} \Sigma_{\Theta nn'}, & \text{for} \quad t = t' \quad n \neq n'; \\ 0, & \text{for} \quad t \neq t' \quad n \neq n'. \end{cases} \tag{3.1.3}$$

Exchangeability of squared system evolution residuals in time

We assume that the squared evolution residuals, $\epsilon^2_{\Theta nt}$, are also second order exchangeable in time for each parameter choice. Let:

$$\epsilon^2_{\Theta nt} = [\mathcal{R}_t(\epsilon_{\Theta n})]^2 = V_{nt} \tag{3.1.4}$$

with:

$$E(V_{nt}) = \Sigma_{\Theta n} \quad \text{Var}(V_{nt}) = \Phi_{V_n} \quad \text{Cov}(V_{nt}, V_{nt'}) = \Xi_{V_n} \quad t \neq t'$$

Here Ξ_{V_n} and Φ_{V_n} are fourth moments which need to be specified. From theorem 2.3.2 this leads to representation statements for the squared evolution residuals for every parameter $n = 1, 2, ...N$:

$$V_{nt} = \mathcal{M}(V_n) + \mathcal{R}_t(V_n) \tag{3.1.5}$$

where:

$$E(\mathcal{M}(V_n)) = \Sigma_{\Theta n} \qquad \text{Var}(\mathcal{M}(V_n)) = \Xi_{V_n} \qquad \text{Var}(\mathcal{R}_t(V_n)) = \Phi_{V_n} - \Xi_{V_n}$$

Exchangeability of mean variance across subsets of parameters

There may be circumstances where we can also express relationships of squared residuals across parameters for one or more subsets of parameters. Where appropriate we then assume second order exchangeability for $\mathcal{M}(V_n)$ across parameters, for a subset of $P \leq N$ parameters, indexed over p, such that from theorem 2.3.2:

$$\mathcal{M}(V_p) = W_p = \mathcal{M}(W) + \mathcal{R}_p(W) \tag{3.1.6}$$

where:

$$E(W_p) = \Sigma_W \quad \text{Var}(W_p) = \Phi_W \quad \text{Cov}(W_p, W_{p'}) = \Xi_W \quad p \neq p'$$

so together with the other assumptions we get:

$$E(\epsilon^2_{\Theta pt}) = E(V_{pt}) = E(\mathcal{M}(V_p)) = E(W_p) = E(\mathcal{M}(W)) = \Sigma_W, \quad \text{Var}(\mathcal{M}(V_p)) = \Phi_W$$

and:

$$\text{Var}(\mathcal{M}(W)) = \Xi_W \quad \text{and} \quad \text{Var}(\mathcal{R}_p(W)) = \Phi_W - \Xi_W$$

and for each $\epsilon^2_{\Theta pt}$ we can write:

$$\begin{aligned}\epsilon^2_{\Theta pt} &= \mathcal{M}(V_p) + \mathcal{R}_t(V_p) \\ &= \mathcal{M}(W) + \mathcal{R}_p(W) + \mathcal{R}_t(V_p)\end{aligned}$$

3.2 Local effects model

3.2.1 Local effects model

Here we model the local surface behaviour in more detail. We assume that local effects are specific to particular components and that all relationships between components are modelled through the global effects model. Local effects are modelled by r_{lct}; the local effects at location l, on component c, at time t:

$$r_{lct} = g(r_{lc(t-1)}) + \epsilon_{rlct} \tag{3.2.7}$$

where the function g is some function which describes the local surface. ϵ_{rlct} is the local effects evolution residual, at location l, on component c, at time t. The properties of the local effects residual are given in section 3.2.2.

3.2.2 Exchangeable residual structures

We use exchangeability assumptions to describe beliefs about the local effects evolution residual. Exchangeability is discussed in section 2.3. We assume that ϵ_{rlct} is second order exchangeable over time, location and components with:

$$E(\epsilon_{rlct}) = 0 \quad \mathrm{Var}(\epsilon_{rlct}) = \Sigma_r \quad \mathrm{Cov}(\epsilon_{rlct}, \epsilon_{rl'c't'}) = 0 \quad \forall \quad l \neq l' \mathrm{or} c \neq c' \mathrm{or} t \neq t'$$

So by theorem 2.3.2 we get:

$$\epsilon_{rlct} = \mathcal{M}(\epsilon_r) + \mathcal{R}_{lct}(\epsilon_r) \tag{3.2.8}$$

and in this case $\mathcal{M}(\epsilon_r) = 0$ so this reduces to:

$$\epsilon_{rlct} = \mathcal{R}_{lct}(\epsilon_r) \tag{3.2.9}$$

and we get that:

$$\mathrm{Var}(\mathcal{R}_{lct}(\epsilon_r)) = \Sigma_r.$$

3.3 System state

3.3.1 True system state

Given global and local models in equations 3.1.1 and 3.2.7 respectively, we model the $C \times 1$ vector of the true system state, Z_{lt} at location l, and time t, by:

$$Z_{lct} = (F\Theta_t)_c + r_{lct} \tag{3.3.10}$$

where F is a known, $C \times N$, dynamic regression matrix, Θ_t is, $N \times 1$, vector of parameter values from the global effects model (equation 3.1.1), then $(F\Theta_t)_c$ is the global effects for component c, and r_{lct} is the local effects term.

3.3.2 Observations of the model

We now model the observation process of the system. The measurement process produces information that summarises the current state of a component. Observations, Y_{ct} for component c, at time t, are made subject to error, ϵ_{Ylct} at each location l, for component c and time t. We model the observation process as:

$$Y_{ct} = f_l\left(Z_{lct} + \epsilon_{Ylct}\right) \tag{3.3.11}$$

where f_l is some function applied over the space of locations, l, which has the property that (from equation 3.3.10):

$$\begin{aligned} Y_{ct} &= f_l\left(Z_{lct} + \epsilon_{Ylct}\right) \\ &= f\left((F\Theta_t)_c + r_{lct} + \epsilon_{Ylct}\right) \\ &= (F^*\Theta_t)_c + f_l\left(r_{lct} + \epsilon_{Ylct}\right) \end{aligned} \tag{3.3.12}$$

where, F^*, is a $C \times N$ dynamic regression matrix. The function, f_l then preserves the global modelling term Θ_t. Possible measurement processes covered by the function f_l, include: averages, quantiles, minima, maxima, modes and medians.

Example: mean function

Let f_l be the mean over locations. From equation 3.3.12:

$$\begin{aligned} Y_{ct} &= f_l\left(Z_{lct} + \epsilon_{Ylct}\right) \\ &= \frac{1}{L_c}\sum_{l=1}^{L_c}\left((F\Theta_t)_c + r_{lct} + \epsilon_{Ylct}\right) \\ &= (F\Theta_t)_c + \frac{1}{L_c}\sum_{l=1}^{L_c}\left(r_{lct} + \epsilon_{Ylct}\right) \end{aligned}$$

where, L_c, is the total number of locations for component, c. In this case:

$$(F^*\Theta_t)_c = (F\Theta_t)_c$$

3.3.3 Exchangeable residual structures

We make the assumption that the measurement error is second order exchangeable with respect to time, components and locations where:

$$E(\epsilon_{Ylct}) = 0 \quad \text{Var}(\epsilon_{Ylct}) = \Sigma_Y \quad \text{Cov}(\epsilon_{Ylct}, \epsilon_{Yl't'c'}) = 0 \quad \forall \quad l \neq l' \quad \text{or} \quad c \neq c' \quad \text{or} \quad t \neq t'$$

using the representation theorem 2.3.2 this leads to:

$$\epsilon_{Ylct} = \mathcal{R}_{lct}(\epsilon_Y)$$

3.3.4 Complete model

From equations 3.1.1, 3.2.7, 3.3.10 and 3.3.11 we then get:

$$\begin{aligned} \text{Observation Equation:} \quad & Y_{ct} = f_l\left(Z_{lct} + \epsilon_{Ylct}\right) \\ \text{True System State:} \quad & Z_{lct} = (F\Theta_t)_c + r_{lct} \\ \text{Global Effects Model:} \quad & \Theta_t = G\Theta_{t-1} + \epsilon_{\Theta t} \\ \text{Local Effects Model:} \quad & r_{lct} = g(r_{lc(t-1)}) + \epsilon_{rlct} \end{aligned}$$

3.4 Discussions of exchangeability assumptions

Second order exchangeability assumptions are weaker than the typical independence type assumptions often made in DLM's. A clear advantage of exchangeability assumption across components is the ability to exploit the potentially large numbers

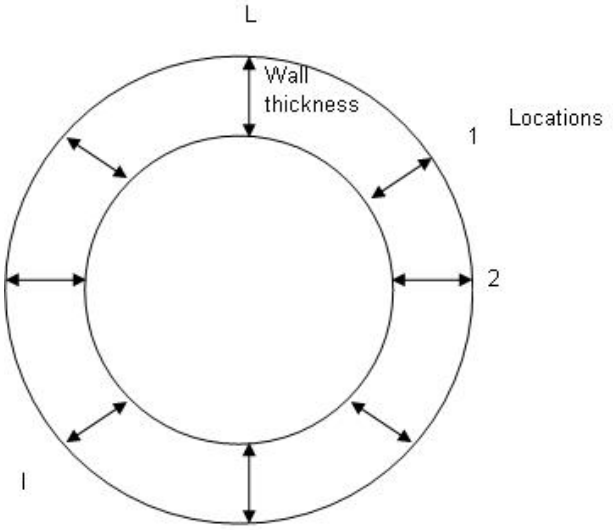

Figure 3.1: Cross Section of a pipework component. Wall thickness is shown as the distance between the inner and outer walls of the pipe. This is recorded by a measurement process across a number of locations and the minimum wall thickness recorded

of components which exist in these types of systems. Specific knowledge about the system characteristics allows a more detailed specification, for example partitioning components into subsets of exchangeable components. Together with the Bayes linear approach, second order exchangeability also gives us the ability to update our beliefs.

3.5 Example: corrosion model

We illustrate the model using the offshore platform application introduced in section 1.4.1. Inspection measurements are used to characterise the extent of corrosion within the system. Different measuring devices including ultrasonic and radiographic are available. Typically, the inspection device generates values for the minimum wall thickness (or maximum pit depth) corresponding to the surface area inspected (the "inspection footprint"). Figure 3.1 shows a diagram of the wall thickness of component, a section of pipework, recorded across L locations. The minimum wall thickness over all locations is recorded.

3.5.1 Global effects model

For the global corrosion model we model the couplet of parameters:

$$\Theta_t = \begin{pmatrix} X_t \\ \alpha_t \end{pmatrix} \tag{3.5.13}$$

where X_t is the $C \times 1$, vector of "global" wall thicknesses at time t (with components X_{ct}), and α_t is the $C \times 1$ vector of corrosion rates (with components α_{ct}) for C components. The global effects model from equation 3.1.1, is then given by:

$$\Theta_t = \begin{pmatrix} I_C & I_C \\ 0_C & I_C \end{pmatrix} \Theta_{t-1} + \epsilon_{\Theta t} \tag{3.5.14}$$

where I_C is the, $C \times C$, identity matrix and 0_C is the, $C \times C$, matrix of zeros. In this example:

$$\begin{aligned} G &= \begin{pmatrix} I_C & I_C \\ 0_C & I_C \end{pmatrix} \\ \epsilon_{\Theta t} &= \begin{pmatrix} \epsilon_{Xt} + \epsilon_{\alpha t} \\ \epsilon_{\alpha t} \end{pmatrix} \\ \Sigma_\Theta &= \begin{pmatrix} \Sigma_X + \Sigma_\alpha & \Sigma_\alpha \\ \Sigma_\alpha & \Sigma_\alpha \end{pmatrix} \end{aligned} \tag{3.5.15}$$

This is a linear growth DLM for wall thickness, X_{ct} and corrosion rate, α_{ct}, both of which evolve in time according to the random evolution errors ϵ_{Xct} and $\epsilon_{\alpha ct}$. Alternatively, equation 3.5.15 can be written:

$$\begin{aligned} X_{ct} &= X_{c(t-1)} + \alpha_{ct} + \epsilon_{Xct} \\ \alpha_{ct} &= \alpha_{c(t-1)} + \epsilon_{\alpha ct} \end{aligned} \tag{3.5.16}$$

Second order exchangeability of ϵ_{Xct}

We have that:

$$\epsilon_{\Theta t} = \begin{pmatrix} \epsilon_{Xt} + \epsilon_{\alpha t} \\ \epsilon_{\alpha t} \end{pmatrix}$$

We assume that ϵ_{Xct} are second order exchangeable in time and correlated across components. i.e.:

$$E(\epsilon_{Xct}) = 0 \quad \text{Var}(\epsilon_{Xct}) = \Sigma_{Xc} \quad \text{Cov}(\epsilon_{Xct}, \epsilon_{Xct'}) = 0 \quad t \neq t'$$

So by theorem 2.3.2 we get:

$$\epsilon_{Xct} = \mathcal{R}_t(\epsilon_{Xc})$$

We also assume that for fixed time t, the residuals ϵ_{Xct} are related across components such that:

$$\text{Cov}[\epsilon_{Xct}, \epsilon_{Xc't'}] = \begin{cases} \Sigma_{Xcc'}, & \text{for} \quad t = t' \quad c \neq c'; \\ 0, & \text{for} \quad t \neq t' \quad c \neq c'. \end{cases} \tag{3.5.17}$$

We also assume that the squared residuals, ϵ^2_{Xct}, are second order exchangeable in time. As in equation 3.1.4 let:

$$\epsilon^2_{Xct} = [\mathcal{R}_t(\epsilon_{Xc})]^2 = V_{Xct}$$

From theorem 2.3.2 this leads to representation statements for the squared evolution residuals for every component $c = 1, 2, ...C$:

$$V_{Xct} = \mathcal{M}(V_{Xc}) + \mathcal{R}_t(V_{Xc})$$

As in equation 3.1.6 we assume that $\mathcal{M}(V_{Xc})$ are also exchangeable across components such that:

$$\mathcal{M}(V_{Xc}) = W_{Xc} = \mathcal{M}(W_X) + \mathcal{R}_c(W_X)$$

where:

$$E(W_{Xc}) = \Sigma_{W_X} \quad \text{Var}(W_{Xc}) = \Phi_{W_X} \quad \text{Cov}(W_{Xc}, W_{Xc'}) = \Xi_{W_X} \quad c \neq c'$$

so together with the other assumptions we get:

$$E(\epsilon^2_{Xct}) = E(V_{Xct}) = E(\mathcal{M}(V_{Xc})) = E(W_{Xc}) = E(\mathcal{M}(W_X)) = \Sigma_{W_X}$$

and for each ϵ^2_{Xct} we can write:

$$\begin{aligned} \epsilon^2_{Xct} &= \mathcal{M}(V_{Xc}) + \mathcal{R}_t(V_{Xc}) \\ &= \mathcal{M}(W_X) + \mathcal{R}_c(W_X) + \mathcal{R}_t(V_{Xc}) \end{aligned}$$

Second order exchangeability of $\epsilon_{\alpha ct}$

We assume that $\epsilon_{\alpha ct}$ are second order exchangeable in time and correlated across components. i.e.:

$$E(\epsilon_{\alpha ct}) = 0 \quad \text{Var}(\epsilon_{\alpha ct}) = \Sigma_{\alpha c} \quad \text{Cov}(\epsilon_{\alpha ct}, \epsilon_{\alpha ct'}) = 0 \quad t \neq t'$$

So by theorem 2.3.2 we get:

$$\epsilon_{\alpha ct} = \mathcal{R}_t(\epsilon_{\alpha c})$$

We also assume that for fixed time t, the residuals $\epsilon_{\alpha ct}$ are related across components such that:

$$\text{Cov}[\epsilon_{\alpha ct}, \epsilon_{\alpha c't'}] = \begin{cases} \Sigma_{\alpha cc'}, & \text{for} \quad t = t' \quad c \neq c'; \\ 0, & \text{for} \quad t \neq t' \quad c \neq c'. \end{cases} \tag{3.5.18}$$

We also assume that the squared residuals, $\epsilon^2_{\alpha ct}$, are second order exchangeable in time. Let:

$$\epsilon^2_{\alpha ct} = [\mathcal{R}_t(\epsilon_{\alpha c})]^2 = V_{\alpha ct}$$

from theorem 2.3.2 this leads to representation statements for the squared evolution residuals for every component $c = 1, 2, ...C$:

$$V_{\alpha ct} = \mathcal{M}(V_{\alpha c}) + \mathcal{R}_t(V_{\alpha c})$$

we assume that $\mathcal{M}(V_{\alpha c})$ are also exchangeable across components such that:

$$\mathcal{M}(V_{\alpha c}) = W_{\alpha c} = \mathcal{M}(W_\alpha) + \mathcal{R}_c(W_\alpha)$$

where:

$$E(W_{\alpha c}) = \Sigma_{W_\alpha} \quad \text{Var}(W_{\alpha c}) = \Phi_{W_\alpha} \quad \text{Cov}(W_{\alpha c}, W_{\alpha c'}) = \Xi_{W_\alpha} \quad c \neq c'$$

so together with the other assumptions we get:

$$E(\epsilon^2_{\alpha ct}) = E(V_{\alpha ct}) = E(\mathcal{M}(V_{\alpha c})) = E(W_{\alpha c}) = E(\mathcal{M}(W_\alpha)) = \Sigma_{W_\alpha}$$

and for each $\epsilon^2_{\alpha ct}$ we can write:

$$\begin{aligned} \epsilon^2_{\alpha ct} &= \mathcal{M}(V_{\alpha c}) + \mathcal{R}_t(V_{\alpha c}) \\ &= \mathcal{M}(W_\alpha) + \mathcal{R}_c(W_\alpha) + \mathcal{R}_t(V_{\alpha c}) \end{aligned}$$

Therefore, $\text{Var}(\epsilon_{\Theta t})$, using equations, 3.5.15, 3.5.17 and 3.5.18 is given by:

$$\text{Var}(\epsilon_{\Theta t}) = \Sigma_\Theta = \begin{pmatrix} \Sigma_X + \Sigma_\alpha & \Sigma_\alpha \\ \Sigma_\alpha & \Sigma_\alpha \end{pmatrix}$$

3.5.2 Local effects model

The local corrosion model describes how the wall thickness of each component deviates from the mean for that component. We assume that the local corrosion, r_{lct}, for a particular component, c is modelled at a number of locations l within the component at time t and is given by:

$$r_{lct} = r_{lc(t-1)} + \epsilon_{rlct}$$

where ϵ_{rlct} is a location effect evolution residual. We assume second order exchangeability assumptions as in equation 3.2.8. We assume that the local corrosion of each component is independent of other components. The function, g, from equation 3.2.7 is simply the identity function.

3.5.3 System State

The $C \times 1$ vector of wall thickness, Z_{lct}, at time, t, observed at location, l, on the component is given by the sum of the global and local corrosion terms. From equation 3.3.10, X_t and r_{lt} respectively as follows:

$$\begin{aligned} Z_{lt} &= \begin{pmatrix} I_C & 0_C \end{pmatrix} \Theta_t + r_{lt} \\ &= X_t + r_{lt} \end{aligned} \tag{3.5.19}$$

In this example:

$$F = \begin{pmatrix} I_C & 0_C \end{pmatrix}$$

3.5.4 Observation process

Only the minimum value of observations at each location, Y_{ct}, is recorded:

$$Y_{ct} = \min_l \left(Z_{lct} + \epsilon_{Ylct} \right)$$

where, ϵ_{Ylct} is the measurement error. We make second order exchangeability assumptions about ϵ_{Ylct} as in section 3.3.2. This can be broken down using equation

3.5.19 to give:

$$\begin{aligned} Y_{ct} &= \min_{l} \left(Z_{lct} + \epsilon_{Ylct} \right) \\ &= \min_{l} \left(X_{ct} + r_{lct} + \epsilon_{Ylct} \right) \\ &= X_{ct} + \min_{l} \left(r_{lct} + \epsilon_{Ylct} \right) \end{aligned} \tag{3.5.20}$$

In this case the function f_l is the minimum over the set of locations.

Examples using the corrosion model will follow in subsequent chapters in sections, 1.4.1, 3.5, 4.9, 5.7.1, 5.7.2, 6.4, 7.7 and 7.9.4).

Chapter 4

Adjusting beliefs about mixed linear temporal models using Bayes linear adjustment and simulation

In chapter 3 we considered the problem of modelling the evolution of complex systems degrading in time. In this chapter we consider using historical data to update our knowledge of the system. We show how to use Bayes linear analysis, introduced in section 2.2, incorporating historical data and combining with a partial prior specification to learn as much as we can about the current state of the system. When modelling large systems, it can be difficult or impractical to make full prior belief specifications. Bayes linear analysis is able to accommodate large systems whilst avoiding too many non-physical simplifications. Updating beliefs is computationally efficient for high dimensional problems. The use of a Bayes Linear approach becomes particularly important when evaluating inspection designs as discussed in chapter 8. A very large space of potential inspection designs requires the need for methods to update beliefs quickly and allow time to evaluate as many inspection designs as possible.

In this chapter we consider the process of updating our beliefs using adjusted expectations. We examine issues involved in the elicitation of beliefs about the

model and the specification of a partial belief structure. We then consider the use of simulation of the model to compute empirical estimates of useful quantities needed to carry our updating and diagnostic checks of the prior specification and adjusted expectations.

4.1 Adjusting beliefs using observational data

Consider a system with C components, where we make observations, Y, of all the components for time 1 to T:

$$Y = (Y_{11}, \ldots, Y_{ct}, \ldots, Y_{CT})^T$$

We then want to use the observational data to update our beliefs about the current state of the system. We compute adjusted beliefs, combining prior beliefs about the system state and the observational data.

Given a collection of observations Y, to update beliefs about the parameter space, Θ_T at time T, we compute the adjusted expectation, $E_Y(\Theta_T)$:

$$\begin{aligned} E_Y(\Theta_T) &= E(\Theta_T) + \mathrm{Cov}(\Theta_T, Y)[\mathrm{Var}(Y)]^{-1}(Y - E(Y)) \\ \mathrm{Var}_Y(\Theta_T) &= \mathrm{Var}(\Theta_T) - \mathrm{Cov}(\Theta_T, Y)[\mathrm{Var}(Y)]^{-1}\mathrm{Cov}(Y, \Theta_T) \end{aligned} \tag{4.1.1}$$

or alternatively if we want to update out beliefs about the true system state, at time T, Z_T, we compute the adjusted expectation, $E_Y(Z)$:

$$\begin{aligned} E_Y(Z_T) &= E(Z_T) + \mathrm{Cov}(Z_T, Y)[\mathrm{Var}(Y)]^{-1}(Y - E(Y)) \\ \mathrm{Var}_Y(Z_T) &= \mathrm{Var}(Z_T) - \mathrm{Cov}(Z_T, Y)[\mathrm{Var}(Y)]^{-1}\mathrm{Cov}(Y, Z_T) \end{aligned}$$

For each of these cases, to be able to update beliefs we need to be able to specify prior beliefs, namely, expectations, variances and covariances about the quantities, Y, Θ_T and Z_T.

4.2 Specifying beliefs

Meaningful prior belief specification for large problems is usually very difficult. Even in small problems, with few sources of uncertainty, it can be difficult to assess a

satisfactory full joint prior probability specification over all of the possible outcomes. In practical problems such as modelling complex industrial systems, there may be hundreds of relevant sources of uncertainty about which prior judgements are made. In such problems, it is arguably impossible for us to carry out a full Bayesian analysis. If such a full prior specification were possible, it would often be the case that the specification was too time consuming and too difficult to check.

To begin specifying beliefs, we must first identify those quantities for which beliefs need to be made and assess sources of uncertainty. The use of expert judgement and auxiliary datasets helps us to specify beliefs. When updating beliefs using Bayes linear analysis only a partial belief specification of expectations, variances and covariances need to be made.

To specify beliefs about the system, we must consider several different aspects. These are discussed in the following sections. Specification of model form is discussed in section 4.3 using the model set-up in chapter 3. Specification of partial belief structures, starting conditions for the model, (co)variance structures and uncertainties in the model are discussed in section 4.4. In section 4.5, we look at expressing the link between quantities within the model through the correlation matrices. In section 4.6 we look at how simulations of the model be can used to estimate relevant expectations variances and covariances. Then we discuss methods of carrying out diagnostic checks of the prior specification in section 4.7.

4.3 Model specification

We must decide how to represent quantities and identify which are linked. To do this we must decide upon a specific model structure. Using the general model form described in chapter 3 and summarised in section 3.3.4 we must specify the quantities f_l, F, G and g. The use of graphical models can help to express this structure and the use of exchangeability can help formalise relationships between quantities to aid specification. The example in section 3.5, considers the case of specifying a model for corrosion on an offshore platform.

4.4 Partial belief specification: general considerations

To perform Bayes linear adjustment, a partial prior belief specification needs to be made. We need first and second order moments (means, variance and covariances), about the quantities which we are interested in adjusting and the data we observe. These moments can be specified given distributional assumptions or expert judgements. Alternatively we can simulate realisations of the model to generate empirical estimates for any quantiles of interest. The problem of specifying partial beliefs structures is considered in Goldstein and Wooff [2007, pg 42-57].

4.4.1 Initial conditions

To generate realisations of our model, some starting conditions e.g. initial component integrity and expected component life need to be specified. For a large industrial system, we might suppose that installed equipment has manufacturer specifications regarding its initial state which could be used for this purpose. Expert judgement may also be used to specify starting conditions.

4.4.2 Quantifying uncertainties

We also need to specify the variances/uncertainties within the model. Experts may be prepared to give tolerances or uncertainties about measurement devices as this is a concept about which they are familiar. In these cases it is therefore natural to use these tolerances to get a specification of model variances for these types of errors. In most cases experts will find it harder and be less prepared to specify beliefs about other uncertainties in the model. Specification of bounds or quantiles can help to judge variance, e.g. specifying one and two standard deviation intervals for a quantity would be enough to give an assessment of variance. We could, instead, use auxiliary data from either a related dataset or a partition of the data set to get parameter estimates.

We may find it easier to specify model variances if we separate contributions into

sums of uncorrelated quantities about which we find it easier to express beliefs. A partition of the variance/covariance into uncorrelated terms is similar to expressing limited exchangeability assumptions. Farrow [2003] considers the problem of specifying covariance structures for large complicated systems.

4.4.3 Specification of higher order moments

In certain cases where we consider the relationships between variances in the model, 4th order beliefs or kurtosis also need to be specified. This could be more difficult for experts to conceptualise than variances. Specifications of uncertainties or bounds on variance specifications can be used to give some estimates of these 4th order quantities. Alternatively we could make a distributional assumption from a family using a 1st and 2nd order description. This would fix the 4th moment through the distributional form. For example if $X \sim N(\mu, \sigma^2)$ then the 4th moment is given by:

$$E(X^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$$

4.5 Specifying Correlation Matrices

The correlation structures between parameters within the model from chapter 3 are important since they control the way information is passed within the system. For large complex systems with many components, correlation matrices may be very large. Consequently specifying the correlations for all components pairwise can be impractical if not impossible. We therefore need ways of simplifying the specification, reducing the number of quantities which need to be specified. Exchangeability assumptions allows us to express relationships between sets of quantities and aid the specification of these correlation structures.

4.5.1 (Co-)Variance structures using exchangeability

Using the global effects model equation 3.1.1, section 3.1:

$$\text{Global Effects:} \qquad \Theta_t = G\Theta_{(t-1)} + \epsilon_{\Theta t}$$

where $\text{Var}(\epsilon_{\Theta nt}) = \Sigma_{\Theta n}$ and from equation 3.1.3, $\text{Cov}[\epsilon_{\Theta nt}, \epsilon_{\Theta n't'}] = \Sigma_{\Theta nn'}$.

If we were to assume second order (s.o.) exchangeability of $\epsilon_{\Theta nt}$ over n and t:

$$\epsilon_{\Theta ct} = \mathcal{M}_\Theta + \mathcal{R}_{\Theta nt} \tag{4.5.2}$$

Then $\Sigma_{\Theta n} = \text{Var}(\epsilon_{\Theta nt}) = \Sigma_\Theta$, for all n, and similarly $\Sigma_{\Theta nn'} = \text{Cov}(\epsilon_{\Theta n't'}, \epsilon_{\Theta nt}) = \Gamma_\Theta$, for all n, n', t and t', $n' \neq n$ and $t' \neq t$. So using an exchangeability assumption we get a simple two parameter form for $\Sigma_\Theta = \Sigma_\Theta(\sigma^2_\Theta, \gamma_\Theta)$. More generally, we can consider subsets of parameters about which we specify exchangeability assumptions, leading to a block form for the correlation matrix.

4.5.2 (Co-)Variance structures using adjacency

Alternatively we could specify (co)-variance using a distance based approach.

We define adjacency matrix, S based on the number of intervening components. This adjacency form defines a distance metric. We use the adjacency matrix, S, to define a covariance matrix Σ_X where the covariance decays exponentially with distance. The decay term, ν controls the rate of decay of the correlation with respect to distance:

$$\Sigma_{Xcc'} = \sigma_{Xc}\sigma_{Xc'}e^{-\nu S_{cc'}} \tag{4.5.3}$$

where $S_{cc'}$ is the number of intervening components and ν is the exponential decay rate. In some cases we may have more detailed information regard the exact positions of components. This information could be used to give a more accurate distance matrix. This gives us a method to specify covariance matrices for large numbers of components, if we think the distance is related to the strength of relationships between components.

4.5.3 Example: specification of correlation structures for the corrosion model

Figure 4.1 shows the schematic for a corrosion circuit. Using this corrosion circuit we can generate an adjacency matrix, S, and different correlation matrices, Σ_X, based on the choice of ν. Figure 4.2 shows an adjacency matrix (based on the number of

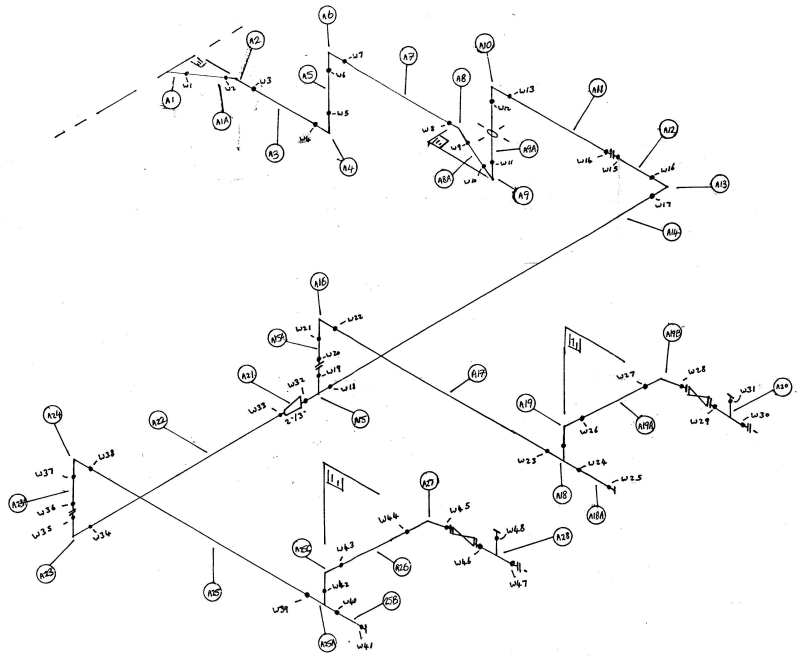

Figure 4.1: Corrosion circuit schematic diagram. Lines shows the route of the pipe. Numbered circles and dots shows position of components.

intervening components) and correlation matrices for different choices of the decay parameter, $\nu = 1$, 0.1, and 0.01.

We can also specify a correlation matrix modelling multiple corrosion circuits. In this case we use a combination of block correlation structures for different corrosion circuits and distance based specification within corrosion circuits. The covariance structure of the wall thickness evolution error, Σ_X, is assumed to take the form of a linear combination of three uncorrelated terms:

- a universal underlying correlation, ρ_0, between all pairs of components (regardless of the circuit(s) to which they correspond).

- a circuit term $\rho_{\text{Cir}}\delta_{\text{CirCir}'}$ where Cir, is the circuit containing component, c, and Cir$'$ is the circuit containing component, c':

$$\delta_{\text{CirCir}'} = \begin{cases} 1 & \text{if Cir} = \text{Cir}' \\ 0 & \text{if Cir} \neq \text{Cir}' \end{cases}$$

.

- a covariance term using the adjacency matrix which decays exponentially at

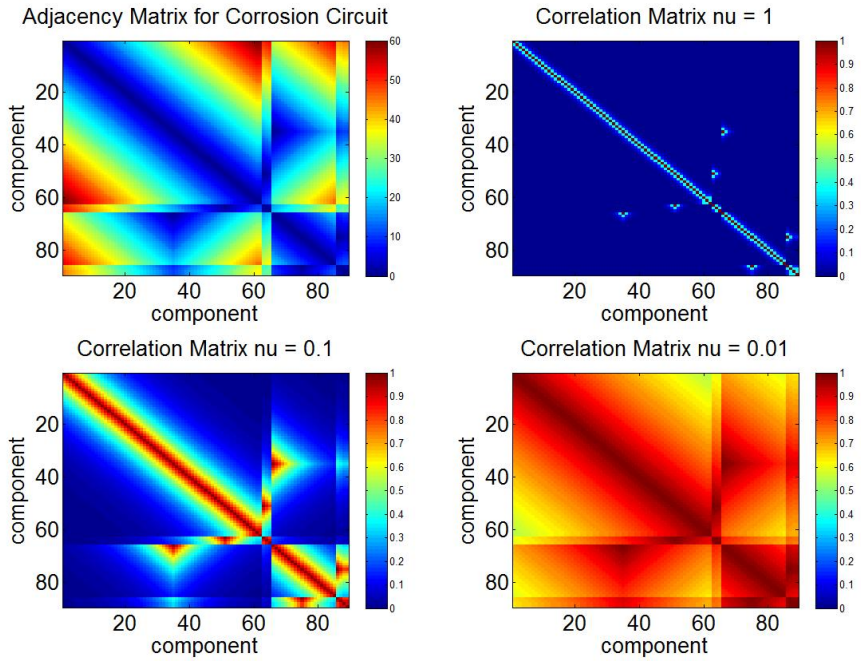

Figure 4.2: Adjacency matrix and corresponding correlation matrices for different values of ν

rate ν with distance S (measured in terms of the number of intervening components along the circuit between the components).

The covariance between components c and c' is given thus:

$$\Sigma_{Xcc'} = \sigma_{Xc}\sigma_{Xc'}\left(\rho_0 + (1-\rho_0)\rho_{\text{Cir}}\delta_{\text{CirCir}'} + (1-\rho_0-(1-\rho_0)\rho_{\text{Cir}})e^{-\nu S_{cc'}}\right) \quad (4.5.4)$$

An example in a case the four corrosion circuits described in section 1.4.1 is illustrated in figure 4.3.

4.6 Simulations of the model

Using ideas in section, 4.2, 4.4 and 4.5 we can make belief statements about what we are prepared to specify. Some quantities we may find hard to specify or find difficult to quantify. Using the limited prior specification Monte Carlo Simulations of the model allow us to get empirical prior expectations, variances and covariances for any quantities of interest. The simulation approach allows us to get an understanding of relationships between all the quantities within the model and give us estimates for

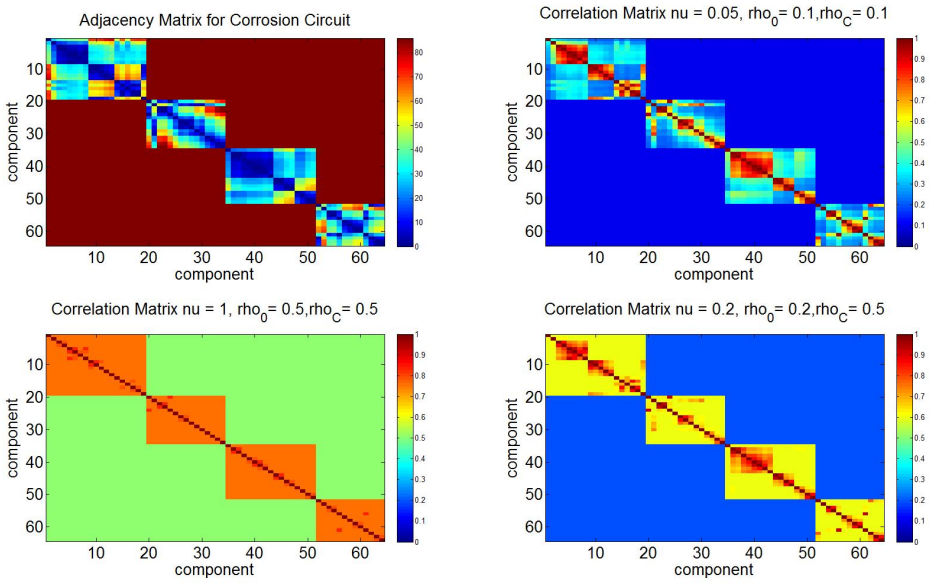

Figure 4.3: Covariance matrix in a case with four corrosion circuits. The four blocks correspond to the intra component distance based correlation in four corrosion circuits

typical model behaviour. Simulation results are used as the basis for prior values used in updating beliefs but also for model forecasting . We can use simulation for modelling uncertainties and check the sensitivity of our prior specification.

Whilst a general DLM can be updated directly, the full corrosion specification and non linear observation equation is too complicated to write down in closed form. However it is straight forward to carry out forward simulation given a partial prior specification. We generate multiple realisations of the model to build up information about relationships between model quantities. This information can then be used to generate empirical assessments for any quantity of interest. The use of simulations means that we can avoid unrealistic simplifying assumptions and modern computing power makes large scale simulations a much more practical proposition. It must be remembered, however, that these simulations do not represent the full uncertainty about the actual system.

The advantage of Bayes linear methods is that only first and second order moments need to be simulated, greatly reducing the computational burden (we may also need to specify fourth order moments if we wish to learn about variances as

well).

4.7 Diagnostic checks

Model diagnostics are an important part of checking whether the outputs from our model are plausible and coherent. We do checks on both our prior specification and our adjusted expectation.

4.7.1 Coherence

The prior specification is coherent if:

$$\mathrm{Var}\begin{pmatrix} Y \\ \Theta_T \end{pmatrix} = \begin{pmatrix} \mathrm{Var}(Y) & \mathrm{Cov}(Y, \Theta_T) \\ \mathrm{Cov}(\Theta_T, Y) & \mathrm{Var}(\Theta_T) \end{pmatrix} \tag{4.7.5}$$

is non-negative definite.

In cases using simulation to get empirical estimates of variance and covariance matrices coherence should hold. However in rank degenerate cases or when the matrices are ill-conditioned it is possible to get rounding errors which can cause these conditions to break down.

This matrix, (equation 4.7.5) is non-negative definite if and only if the following three properties hold (Goldstein and Wooff [2007] page 69):

1. $\mathrm{Var}(Y)$ is non-negative definite
2. $\mathrm{Cov}(Y, \Theta_T) \in \mathrm{range}(\mathrm{Var}(Y))$ i.e. $\mathrm{Cov}(Y, \Theta_T)$ are in the linear span of the columns of $\mathrm{Var}(Y)$.
3. $\mathrm{Var}(\Theta_T) - \mathrm{Cov}(\Theta_T, Y)\mathrm{Var}(Y)^{\dagger}\mathrm{Cov}(Y, \Theta_T)$ is non-negative definite for any generalized inverse of $\mathrm{Var}(Y)$.

4.7.2 Data discrepancy

To check that the prior specification does not conflict with the observed data, we can use the Mahalanobis distance, [Mahalanobis, 1936]. The Mahalanobis distance is a distance measure based on the correlations between variables, [Mahalanobis, 1936].

It differs from the standard Euclidean distance metric in that uses the full covariance matrix so is scale invariant. The Mahalanobis distance or data discrepancy, $\text{Dis}(Y)$, is given by:

$$\text{Dis}(Y) = (Y - E(Y))^T \text{Var}(Y)^\dagger (Y - E(Y)). \tag{4.7.6}$$

Where $\text{Var}(Y)^\dagger$ is the inverse of $\text{Var}(Y)$ if invertible and if not a generalised inverse, such as the Moore-Penrose pseudo inverse should be used. Large discrepancies, may suggest poor specification of $E(Y)$ or that $\text{Var}(Y)$ has been underestimated. Very small discrepancies could suggest that $\text{Var}(Y)$ is too large.

The discrepancy has expected value:

$$\begin{aligned} E(\text{Dis}(Y)) &= E\left[(E(Y) - Y)^T \text{Var}(Y)^\dagger (E(Y) - Y)\right] \\ &= E\left[(Y - E(Y))E(Y - E(Y))^T\right] \text{Var}(Y)^\dagger \\ &= \text{trace}(\text{Var}(Y))\text{Var}(Y)^\dagger \\ &= \text{rank}(\text{Var}(Y)). \end{aligned}$$

It is sometimes easier to compare discrepancies if we standardise these values by dividing by the $\text{rank}(\text{Var}(Y))$. We denote this discrepancy ratio for Y, $\text{Dr}(Y)$, as:

$$\text{Dr}(Y) = \frac{\text{Dis}(Y)}{\text{rank}(\text{Var}(Y))}, \tag{4.7.7}$$

which has prior expectation 1.

4.7.3 Variance of Discrepancy

When comparing discrepancies we also would like to know how large we would expect discrepancies to be. Computing the variance of the discrepancy for an arbitrary distribution would require us to express beliefs about all possible 3rd and 4th moments of a distribution. As discussed in section 4.4.3 the specification of higher order moment for large problems is hard. Vysochanskij and Petunin [1980] and Pukelsheim [1994] describe a "3 σ rule" stating that for any uni-modal continuous random quantity, u with standard deviation σ_u:

$$P(|u - E(u)| \leq 3\sigma_u) \geq 0.95.$$

So an observation 3 standard deviations away from the mean can be used as a diagnostic on our uncertainty about a quantity. Therefore, we can then test if, $\mathrm{Dr}(Y)$ from equation 4.7.7:

$$|1 - \mathrm{Dr}(Y)| > 3\sigma_{\mathrm{Dr}Y}, \tag{4.7.8}$$

where:

$$\sigma_{\mathrm{Dr}Y} = \sqrt{\mathrm{Var}(\mathrm{Dr}(Y)|)}$$

and if it is more than 3 standard deviations away, then we may have a potential problem in our prior specification.

4.7.4 Normal Approximation

In the case where Y is multivariate normal:

$$Y \sim N_{\mathrm{rank}(\mathrm{Var}(Y))}(\mu, \Sigma)$$

then by Imhof [1961] the Mahalanobis distance is χ^2 distributed:

$$\mathrm{Dis}(Y) \sim \chi^2_{\mathrm{rank}(\mathrm{Var}(Y))}$$

and so, subject to normality:

$$\mathrm{Var}(\mathrm{Dis}(Y)) = 2(\mathrm{rank}(\mathrm{Var}(Y))).$$

and so, under normality $\mathrm{Dis}(Y)$):

$$3\sigma_{\mathrm{Dr}Y} = 3\sqrt{2\mathrm{rank}(\mathrm{Var}(Y))} \tag{4.7.9}$$

and so for the discrepancy ratio $\mathrm{Dr}(Y)$), from equation 4.7.7:

$$\mathrm{Var}(\mathrm{Dr}(Y)) = 2\frac{\mathrm{rank}(\mathrm{Var}(Y))}{\mathrm{rank}(\mathrm{Var}(Y))} = 2$$

and so using equation 4.7.9 in the case of the multivariate normal distribution:

$$3\sigma_{\mathrm{Dr}Y} = 3\sqrt{2}$$

Then from equation 4.7.8 we can use:

$$|1 - \mathrm{Dr}(Y)| > 3\sqrt{2}$$

as a diagnostic test to check whether discrepancies are higher than we would expect for the normal distribution.

This can be extended to non-negative definite cases [Liu et al., 2009], but still under the assumption of normality. We could replace the normal distribution by any other location/scale family of specified distributions (e.g. the t distribution) and still extract the higher order moments, using simulation where needed.

4.7.5 Adjustment discrepancy

Using the Mahalanobis distance we can test the adjusted expectations by computing the adjustment discrepancy for each of our updated values:

$$\mathrm{Dis}_Y(\Theta_T) = (E_Y(\Theta_T) - E(\Theta_T))^T \mathrm{RVar}_Y(\Theta_T)^\dagger (E_Y(\Theta_T) - E(\Theta_T)),$$

where $\mathrm{RVar}_Y(\Theta_T)$ is the resolved variance see equation 2.2.3. The adjustment discrepancy ratio is given by:

$$\mathrm{Dr}_Y(\Theta_T) = \frac{\mathrm{Dis}_Y(\Theta_T)}{\mathrm{rank}(\mathrm{RVar}_Y(\Theta_T))}.$$

then test if $|1 - \mathrm{Dr}_Y(\Theta_T)| > 3\sigma_{\mathrm{Dr}Y}$ to check for potential problems. The adjustment discrepancy can be used to see if our adjustment was larger or smaller than expected.

4.7.6 Component-wise discrepancy

We can compute the discrepancy component-wise for each component, c, and time, t:

$$\mathrm{Dis}(Y_{ct}) = \frac{(Y_{ct} - E(Y_{ct}))^2}{\mathrm{Var}(Y_{ct})}.$$

This could identify problems with individual observations rather than overall specification. Hence this approach may help to identify individual outliers. However since this is done point-wise, the covariance structure between the observations is lost, so there may be situations where a single point is considered problematic when in fact it fits perfectly well with the prior specification. In most cases, it is probably preferable to do both since they could identify different types of potential problems.

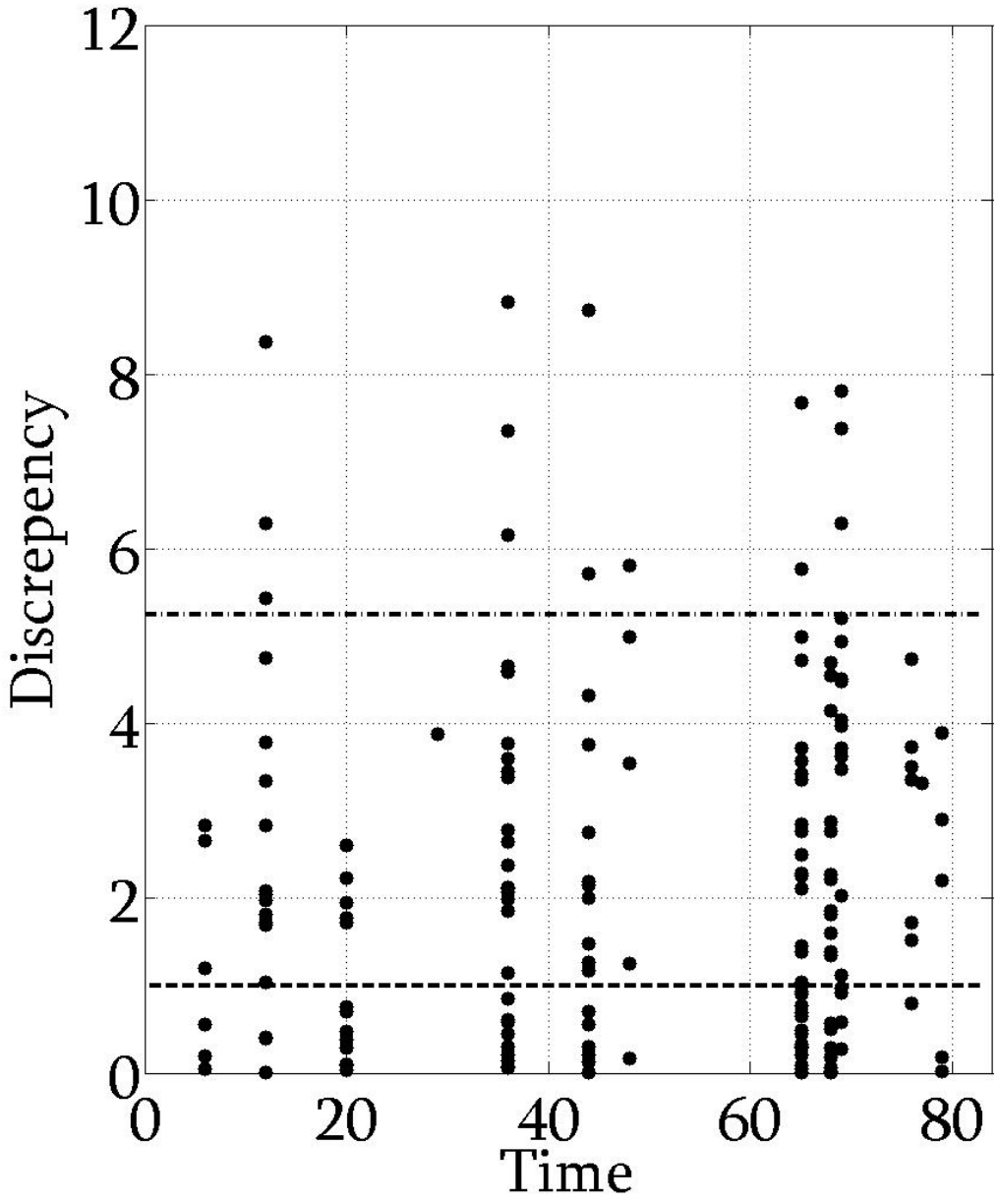

Figure 4.4: Component-wise discrepancy, $\mathrm{Dis}(Y_{ct})$ for a realisation of simulated corrosion data. The expected value of $\mathrm{Dis}(Y_{ct})$ is 1, shown as a horizontal line. Also shown is the horizontal line corresponding to $|1 - \mathrm{Dis}(Y_{ct})| = 3\sqrt{2}$ the $3\sigma_{\mathrm{Dr}Y}$ rule under normality, serving as a warning limit for unusually large values of discrepancy.

Figure 4.4 shows an example of a component-wise discrepancy using simulated data using the corrosion model (for a single component $\text{Dis}(Y_{ct}) = \text{Dr}(Y_{ct})$). Simulated using forward simulation from an initial system system using the model summarised in 3.3.4. In this case 14 of the 175 observations are outside the $|1 - \text{Dis}(Y_{ct})| > 3\sqrt{2}$ test. We would expect 5% of cases to fall outside the 95% bound which would be ≈ 8 so there are more potential outliers than we would expect. If the discrepancy measure shows many value high discrepancies it is a warning that the prior specification and the observed data are incompatible. In these cases further investigation is required to determine the cause of the problem.

4.8 Algorithm for simulation and Bayes linear updating of beliefs

We now present an algorithm for simulating and carrying out the relevant belief adjustments from the model in section 3.3.4.

1. Specify starting conditions (system state) at time 0, locations, correlations and system variance etc.

2. Generate a large number of random evolution residuals $\epsilon_{\Theta 1}$, ϵ_{Ylc1}, ϵ_{rlc1} for all components, c, and locations, l, at time 1. The number of realisations needed scales with the size of the problem. For these type of problems $>$ 10000 simulations would be most likely be needed. Sampling variances can be used to determine if enough variances have been used.

3. Use simulated residuals to generate a set of realisations of Y_{c1}, Z_{lc1}, Θ_1 and r_{lc1} at time 1 using model described in section 3.3.4.

4. Repeat 2-3 stepping model forward one step at a time to the final time step when we want to forecast time T.

5. Calculate relevant moments for adjustment e.g. $E(Y)$, $\text{Var}(Y)$, $\text{Cov}(Y, \theta_T)$, $E(\theta_T)$ and $\text{Var}(\theta_T)$.

6. Perform diagnostic checks for coherence and computing data discrepancy using prior information to check for consistency of prior beliefs.

7. Compute Bayes linear adjustment of the system state where required (e.g. at some future forecast time) .

8. Perform diagnostic checks, computing adjustment discrepancy.

4.9 Example: Application to inspection of an offshore platform

We now consider prior specification and Bayes linear adjustment of wall thickness in the analysis of inspection data from a full-scale offshore platform. We model corrosion using the model as given in section 3.5. We use historical data as introduced in section 1.4.1. Component minimum wall thickness, obtained during inspection campaigns using non-intrusive ultrasonic measurements for the period 1998 - 2005, are available, data is given in appendix A. Based on the frequency of observations and the requirements for inspection planning, we select a monthly time-increment for modelling; the historical period therefore consists of 83 time points.

We first consider an example using Bayes linear adjustment to update beliefs about the wall thickness of a single component, (component 43 from corrosion circuit C in appendix A), about which we have 6 inspections:

Time (Months)	Observed Wall Thickness (mm)
12	12
36	9.5
44	11
65	9.5
69	8
76	8

4.9.1 Prior Specification

We specify partial beliefs about system state as in section 4.4 and generate realisations of the model to get empirical estimates for the means, variances and covariances needed to compute Bayes linear adjustments.

To specify the prior wall thickness for a component we use the manufacturers nominal wall thickness at installation, $X_0 = 13.49$. The prior corrosion rate is specified using auxiliary data as $\alpha_0 = -0.15$.

The prior variances for the model are specified using expert judgement:

$$\Sigma_\alpha = 0.01^2 \qquad \Sigma_X = 0.2^2 \qquad \Sigma_r = 0.2^2 \qquad \Sigma_Y = 0.2^2$$

We also have to specify the size of the grid for the local effects surface. We set the number of locations as 20. We can then use this partial prior specification together with the model and generate 10000 realisations of the model as discussed in section 4.6.

4.9.2 Updating Beliefs

We compute adjusted expectation and variance of the true minimum wall thickness, Z_t using the vector of observations Y, where:

$$Y = \{Y_{12}, Y_{36}, Y_{44}, Y_{65}, Y_{69}, Y_{76}\}.$$

Then we find adjusted beliefs (as in section 4.1) about the wall thickness at all time points and forecast into the future up to time 120:

$$E_Y(Z_t) = E(Z_t) + \mathrm{Cov}(Z_t, Y)[\mathrm{Var}(Y)]^{-1}(Y - E(Y))$$
$$\mathrm{Var}_Y(Z_t) = \mathrm{Var}(Z_t) - \mathrm{Cov}(Z_t, Y)[\mathrm{Var}(Y)]^{-1}\mathrm{Cov}(Y, Z_T)$$

where $E(Z_t)$, $E(Y)$, $\mathrm{Cov}(Z_t, Y)$, $\mathrm{Var}(Y)$ and $\mathrm{Var}(Z_t)$ are all computed using empirical estimation over multiple realisations of the model. Figure 4.5 shows a comparison between the prior and adjusted beliefs when updating a single component. The prior beliefs are shown in blue and the result of the simulation of the model; extending our partial belief specification to tell us about the whole system. The inspection points

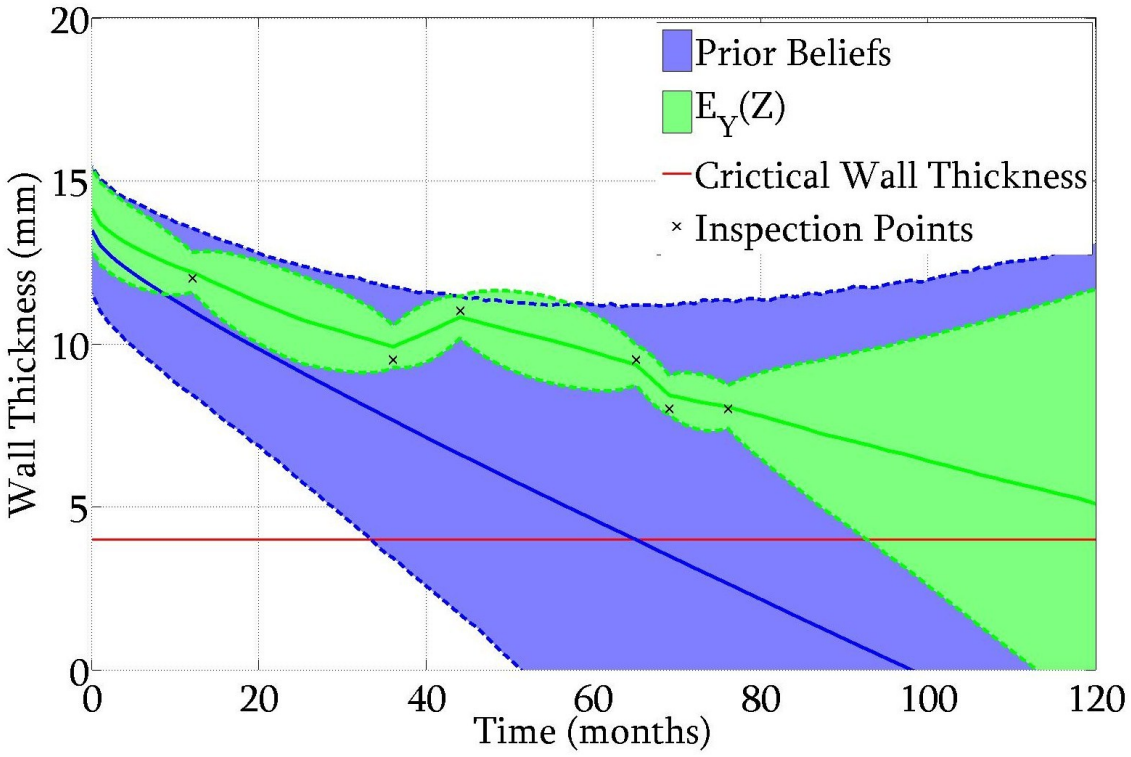

Figure 4.5: Comparison between prior and adjusted beliefs for updating a single component. The prior beliefs are shown in blue and the result of the simulation of the model; extending our partial belief specification to tell us about the whole system. The inspection points are show in black and the adjusted beliefs are show in green. Shaded areas represents 95% uncertainty bounds. Prior uncertainty is derived from simulation and posterior uncertainty is derived from the adjusted variance

are shown in black and the adjusted beliefs are shown in green. We can see that after observing the system values, we were over estimating the rate of degradation of the system. We can also see that, our uncertainty about the system state has decreased and perhaps unsurprisingly we are more certain about the system state close to points which we observed. If we were observing the system without error, the Bayes linear estimate at points for which we observe would be exact. The red line shows a "critical wall thickness" at which point the component would be replaced. After computing adjusted beliefs, it may be discovered that the life span of the component appears longer than previously thought, which would save money replacing a healthy component. Figure 4.6 shows the effect of updating beliefs sequentially for each inspection.

4.9.3 Diagnostics

We can compute the discrepancy on both the data and adjusted beliefs as in section 4.7. The data discrepancy ratio for Y is:

$$\mathrm{Dr}(Y) = 1.703$$

and the adjustment discrepancy is:

$$\mathrm{Dr}_Y(Z) = 1.523$$

where from section 4.7.2 the expected value of $Dr(Z) = 1$ and under an assumption of normality we would not expect the discrepancy ratio to be more than $3\sqrt{2}$, 95% of the time. Figure 4.7 shows the point wise data and adjustment discrepancies. We see that the inspection points at time 44 and 65, are slightly higher than the rest and this is consistent with a comparison between the inspection points with the prior estimate. The inspection points are on the high side. They are not however high enough to be considered outliers.

4.10 Example updating several components

We now consider updating adjusting beliefs about multiple correlated components. (Using data from component 43 and 44 from corrosion circuit C given in appendix

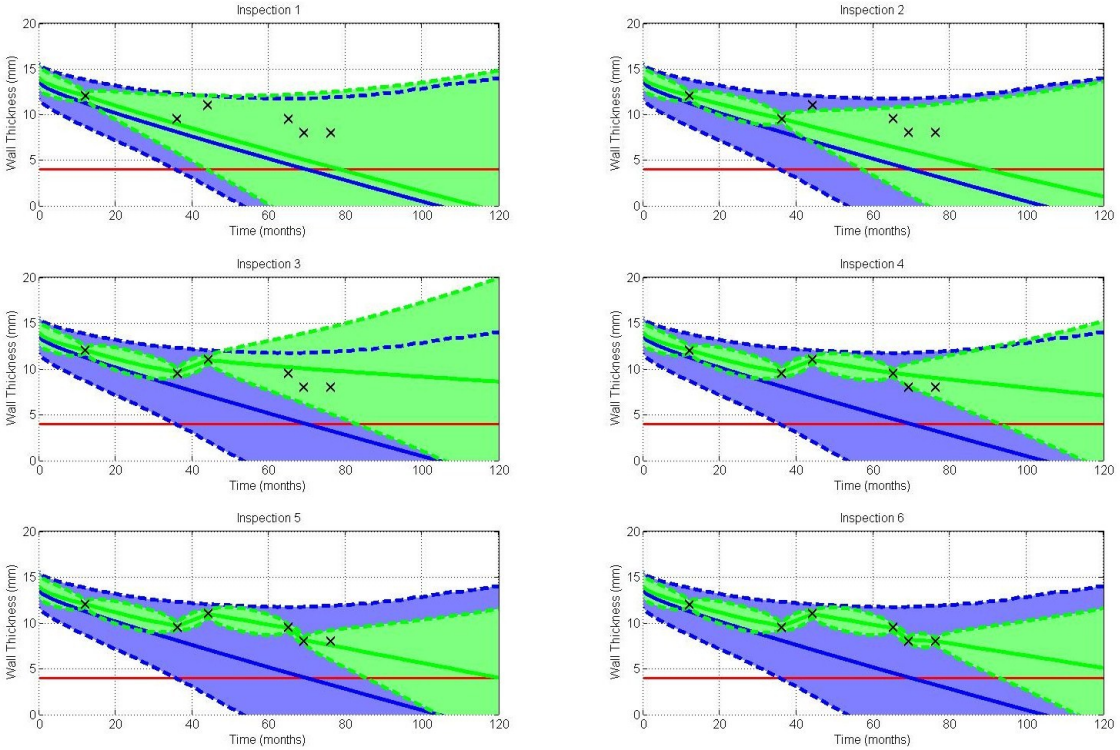

Figure 4.6: Comparison between prior and adjusted beliefs for updating a single component, updating beliefs sequentially after each inspection. The prior beliefs are shown in blue and the result of the simulation of the model; extending our partial belief specification to tell us about the whole system. The inspection points are show in black and the adjusted beliefs are shown in green. Shaded areas represents 95% uncertainty bounds. Prior uncertainty is derived from simulation and posterior uncertainty is derived from the adjusted variance

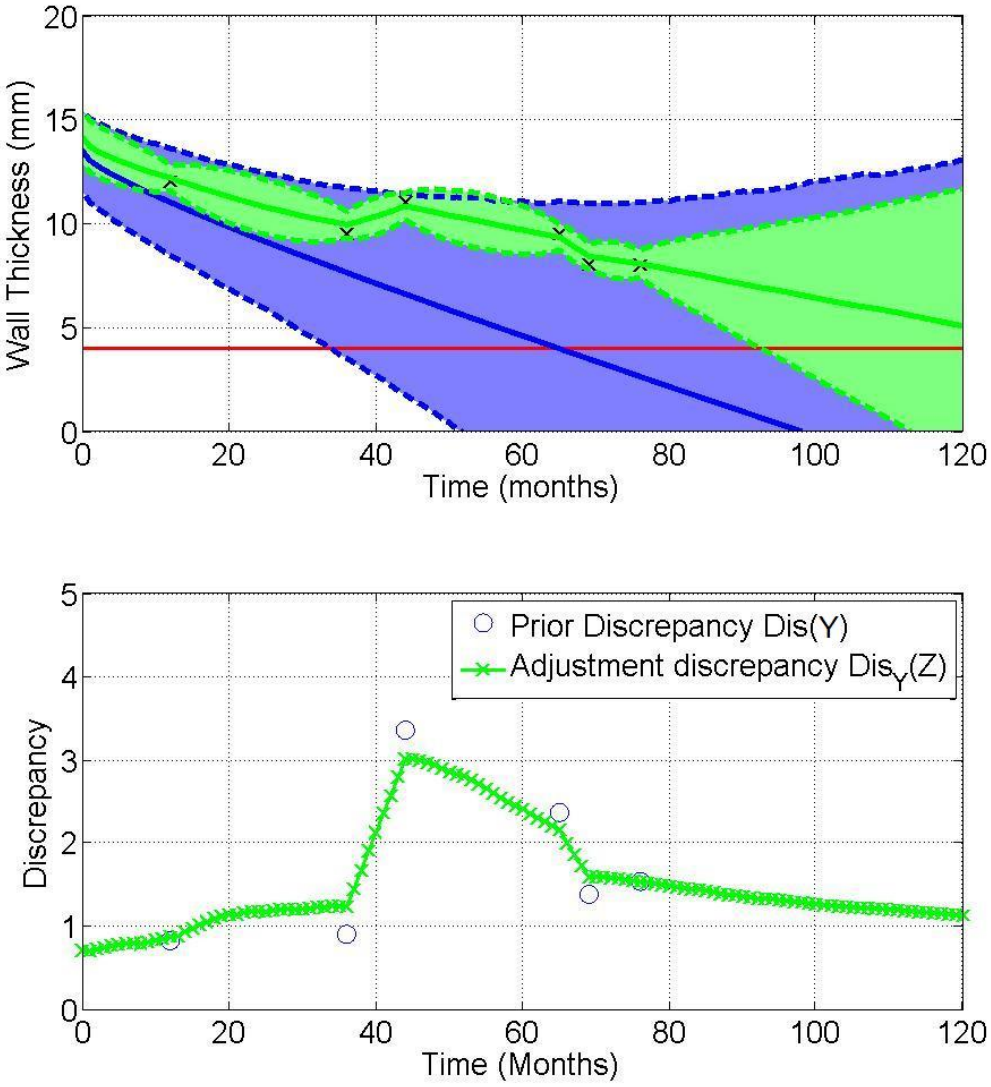

Figure 4.7: Point-wise data discrepancy Dis(Y), blue, and adjustment discrepancy Dis$_Z$(Y), green. The discrepancies are all below the 3σ bound which suggests that there are no obvious outliers. Shaded areas represents 95% uncertainty bounds. Prior uncertainty is derived from simulation and posterior uncertainty is derived from the adjusted variance.

A). We have inspection data for two components and want to update beliefs about these and a third component about which we do not have any observations.

Time (Months)	Observed Wall Thickness (mm)	Component
12	12	43
12	13.49	44
36	9.5	43
44	11	43
65	9.5	43
69	8	43
69	10.5	44
76	8	43
76	10.5	44

Each component has the same prior specification as the single component given in section 4.9.1. We also need to specify the correlation matrices between components as discussed in section 4.5. We specify a flat correlation between all components of $\rho_0 = 0.9$ between all components giving variance matrices;

$$\Sigma_X = \begin{pmatrix} 0.0400 & 0.0360 & 0.0360 \\ 0.0360 & 0.0400 & 0.0360 \\ 0.0360 & 0.0360 & 0.0400 \end{pmatrix},$$

and

$$\Sigma_\alpha = \begin{pmatrix} 0.00010 & 0.00009 & 0.00009 \\ 0.00009 & 0.00010 & 0.00009 \\ 0.00009 & 0.00009 & 0.00010 \end{pmatrix}.$$

Figure 4.8 shows the adjusted beliefs of all three components. Shaded areas represents a 95% uncertainty bounds. We can see that for component 43 and 44 where we inspect we are most certain about the system state. We also learn about the unobserved component even though we have not directly observed it. This learning come through the correlation matrices. For both component 43 and 44 the observed corrosion rate is lower than the prior prediction and consequently we think our life span of the unobserved component is longer too.

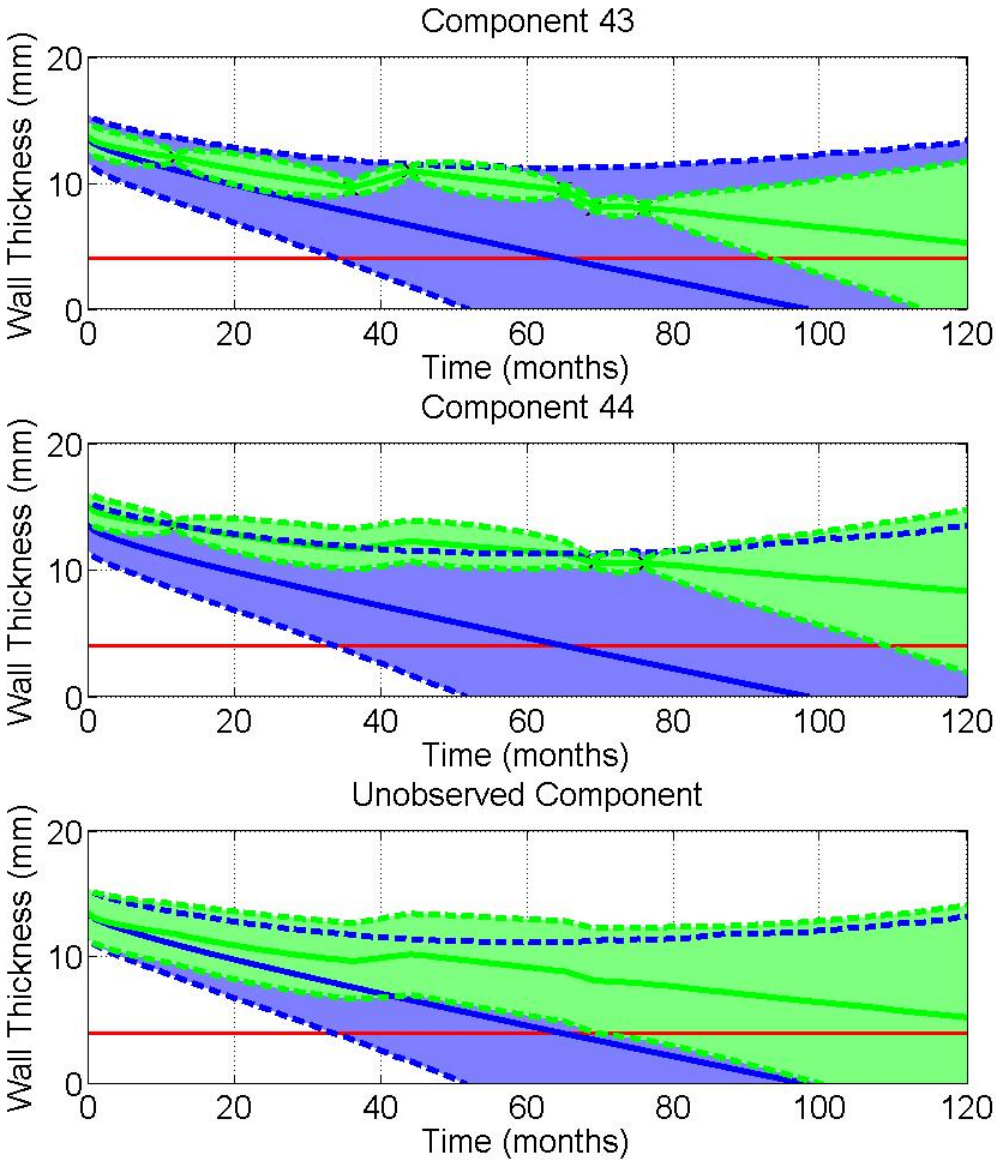

Figure 4.8: Bayes linear adjustment of three components, correlation $\rho_0 = 0.9$. Shaded areas represents a 95% uncertainty bounds. Prior uncertainty is derived from simulation and posterior uncertainty is derived from the adjusted variance. The prior specification for each of the components is the same. Components 43 and 44 show that the adjusted corrosion rate is lower than the prior prediction. We also learn about the unobserved component.

The strength of the correlation between components will control the strength of the adjustment. Figure 4.9 shows the same example with the exception that the correlation between components is weaker; $\rho_0 = 0.3$. Consequently the size of the adjustment from the prior estimates for the unobserved component is reduced. Figure 4.10 shows the discrepancy for the three components. The adjustment discrepancies for components 44 and the unobserved component between times 6 and 56 are above the 3σ diagnostic. However it must be remembered that the point-wise diagnostic loses the covariances between observations and successive time points are not independent.

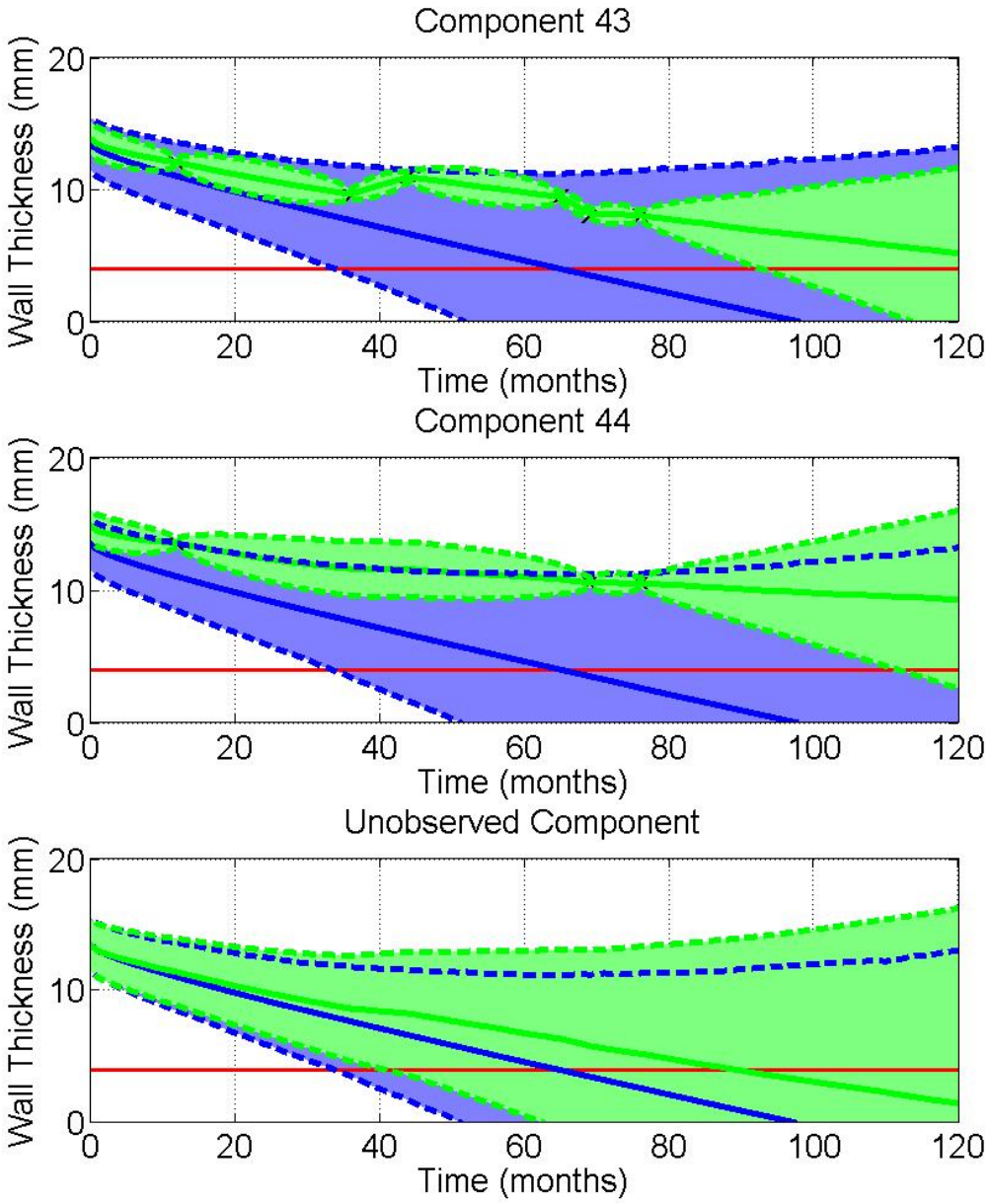

Figure 4.9: Bayes linear adjustment of three components, correlation $\rho_0 = 0.3$. Components 43 and 44 show that the adjusted corrosion rate is lower than the prior prediction. In comparison with figure 4.8, we get a smaller adjustment of beliefs about the unobserved component due to the smaller correlation.

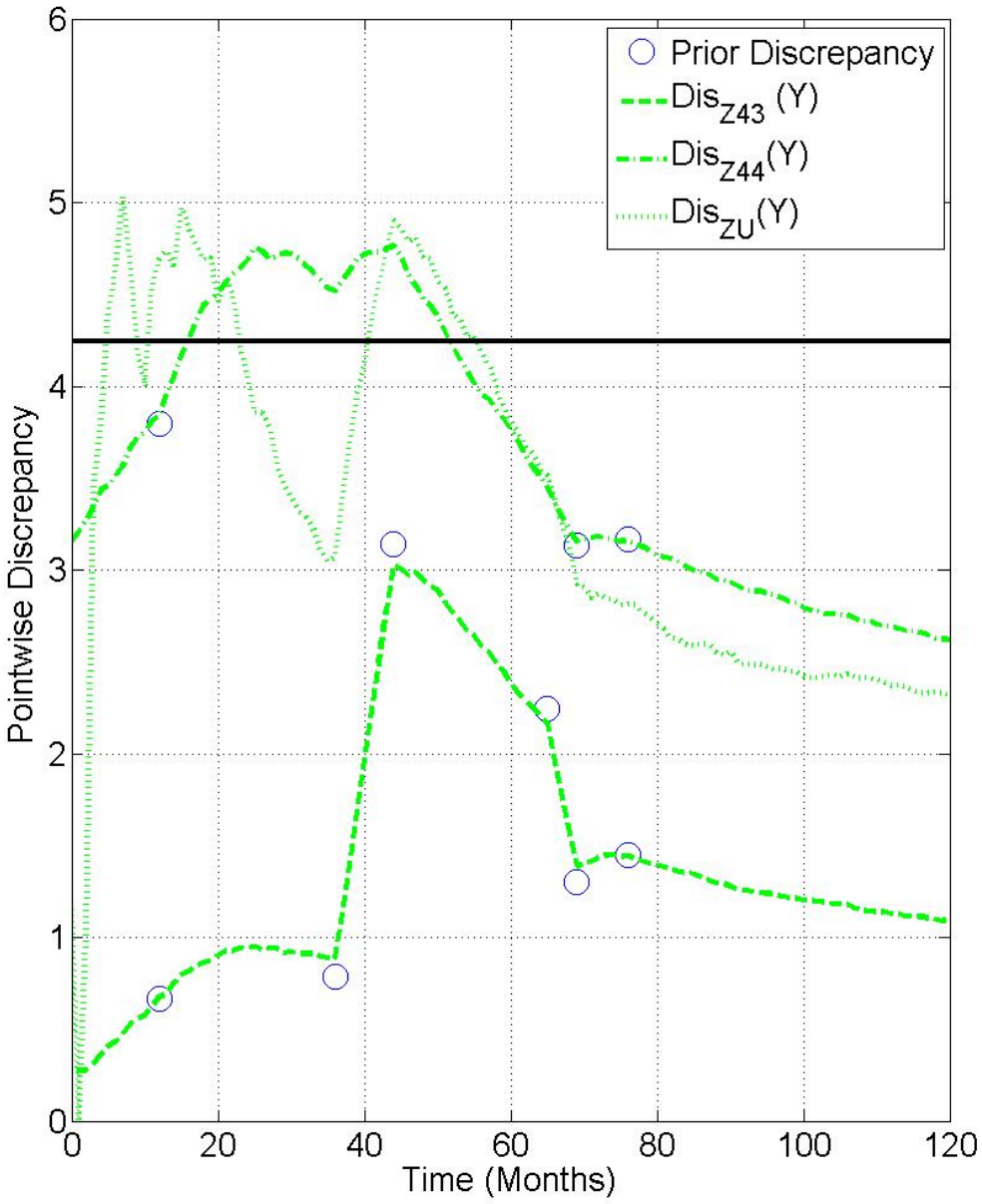

Figure 4.10: Component-wise prior discrepancy Dis(Y), blue, and adjustment discrepancy Dis$_{Z_c}(Y)$, green for the three components. The adjustment discrepancies for components 44 and the unobserved component between times 6 and 56 are above the 3σ diagnostic.

4.11 Comparison of independent and correlated systems

Previous attempts to model corrosion data usually treat components as independent, ignoring any dependency structure. We can explore the effect of this independence assumption on model performance using synthetic data. 50 synthetic data are generated sets using the same prior specification and inspection design as the real historical data given in appendix A. We model each of the 50 data sets, firstly assuming independence and subsequently assuming dependence. We compute adjusted expectations in each case and compare discrepancy diagnostics. Figure 4.11 compares prior and adjustment discrepancy ratios with respect to the independence and dependence model assumptions.

Green and blue lines, respectively, correspond to data and adjustment discrepancy ratios for the independence model; model fit is poor. The extent of adjustment is small since adjustment is component-wise. Yellow and red lines respectively, show prior and adjustment discrepancy ratios for the dependence model; model fit is better. Perhaps unsurprisingly, by fully exploiting correlation structure we share information from all observations across components to learn about the whole system.

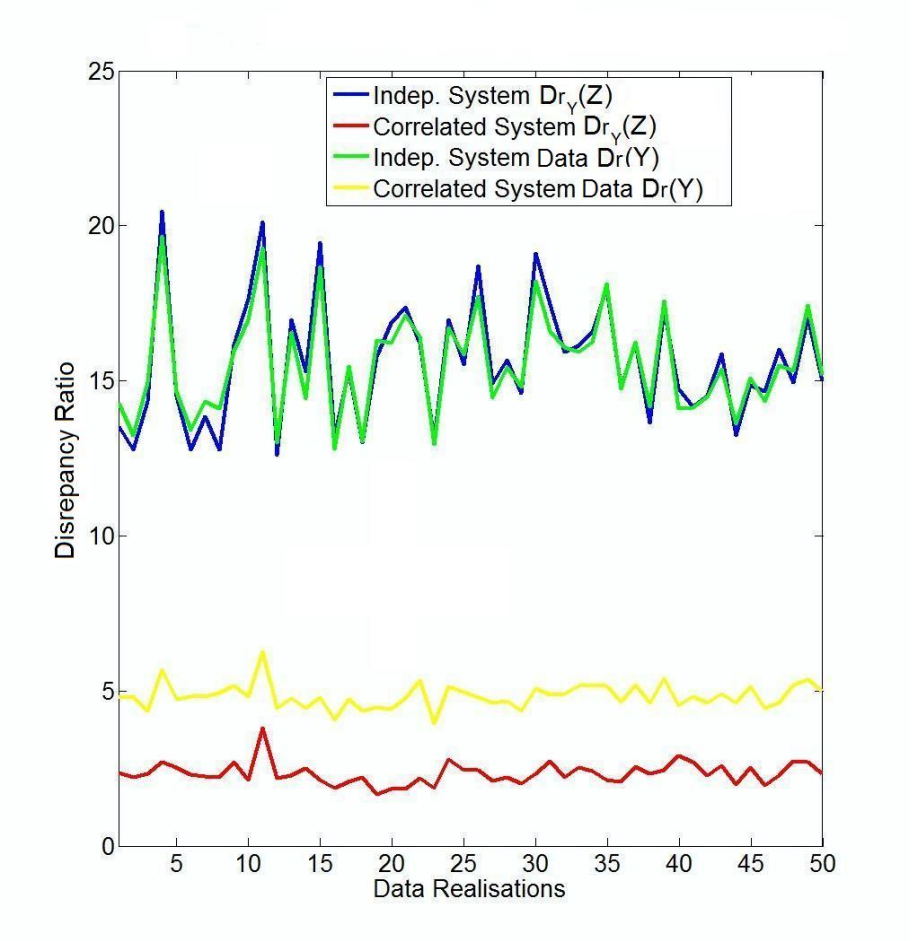

Figure 4.11: Comparison of the discrepancy ratio for cases when system is modelled treating components as independent or correlated for 50 simulated data sets. Green and blue lines respectively, correspond to data and adjustment discrepancy ratios for the independence model; model fit is poor. Yellow and red lines respectively, show data and adjustment discrepancy ratios for the dependence model; model fit is better. By exploiting correlation structure we share information from all observations across components to learn about the whole system.

Chapter 5

Bayes linear variance learning for a linear growth DLM

In the previous chapter, we considered updating our beliefs about the system state using Bayes linear adjustment, the elicitation of prior specifications and the use of simulation to estimate important quantities. However, in Bayes Linear adjustment of our beliefs about the system levels, we are not learning anything about the variance structures within the model. If we don't learn about the variances within the model then we are relying on the prior specification to be accurate, which may well be an unreasonable expectation. We need to construct estimators from the data which are informative for variance learning, whilst dealing with the issues of short time series and irregular inspection data. To be able to learn about variances, we will use a combination of Bayes linear adjustment (section 2.2), and second order exchangeability assumptions (section 2.3).

We will begin by showing how to use Bayes linear adjustment to update the variance of a population with known mean using second order exchangeability assumptions. We will then look at the case of a simple univariate DLM and compute adjusted expectations for the population variance using squared differences of observations in time. Then we consider learning about variances for a multivariate linear growth DLM in the case of complete inspections before examining methods to deal with problems, such as how to deal with irregular inspections, and partial/ incomplete observations. We then look at an example using the oil platform application.

In the case of simple DLM's and the linear growth DLM we could update variance structures using a full Bayes approach relatively easily (section 2.2.1). The intention however, is to introduce and explain the ideas for use in later chapters where a full Bayes approach would prove to be too hard due to the size of the problem, and the complexity of the observation structure.

5.1 Bayes linear learning about population variance

5.1.1 Exchangeability and the representation theorem

Firstly, consider the situation where we wish to learn about the variance of a population with **known mean**, μ_A, [Goldstein and Wooff, 2007, pages 282-283]. We are given a collection of quantities, $A = \{A_1, A_2, \dots\}$, which we assume to be infinitely second order exchangeable (see section 2.3) whose prior specification is:

$$E(A_i) = \mu_A \quad \text{Var}(A_i) = \Sigma_A,$$

We wish to use observational data from this population to update our beliefs about the population variance whose prior expectation is Σ_A. Using the second order representation theorem,(see theorem 2.3.2), we have:

$$A_i = \mathcal{M}(A) + \mathcal{R}_i(A),$$

where since in this case μ_A is known and non–random:

$$\mathcal{M}(A) = \mu_A$$

consequently:

$$\text{Var}(\mathcal{M}(A)) = \text{Cov}(A_i, A_j) = 0, i \neq j.$$

To find a representation which will tell us about the variance we let:

$$V_i = [\mathcal{R}_i(A)]^2 = (A_i - \mu_A)^2$$

We assume that this sequence is also second order exchangeable with prior specification:

$$E(V_i) = \Sigma_A \qquad \text{Var}(V_i) = \Phi_V \qquad \text{Cov}(V_i, V_j) = \Xi_V$$

and hence we can apply the representation theorem for a second time:

$$[\mathcal{R}_i(A)]^2 = V_i = \mathcal{M}(V) + \mathcal{R}_i(V),$$

then:

$$E(\mathcal{M}(V)) = \Sigma_A \qquad \text{Var}(\mathcal{M}(V)) = \Xi_V \qquad \text{Var}(\mathcal{R}_i(V)) = \Phi_V - \Xi_V.$$

The quantities, Φ_V and Ξ_V need to be specified directly as prior beliefs. These are fourth order quantities and so have some relation to the kurtosis or the shape distribution. Specification of higher order moments is discussed in section 4.2.

5.1.2 Population mean

The population mean terms, $\mathcal{M}(V)$, from the representation theorem are the random quantities of interest, whose expectations are informative for variance learning. These terms go into the Bayes linear adjustment equation to update our beliefs.

5.1.3 Beliefs

Given the exchangeability assumptions, we now consider how to update beliefs about the population variance. Therefore, given a collection of n observations, $D = \{A_1 \dots A_n\}$, drawn from the population we can calculate the sample estimate for the variance:

$$\bar{D}_n^{(2)} = \frac{1}{n}\sum_{i=1}^{n}(A_i - \mu_A)^2.$$

we have:

$$\begin{aligned} E(\bar{D}_n^{(2)}) &= E(\frac{1}{n}\sum_{i=1}^{n}(A_i-\mu_A)^2) = E(\frac{1}{n}\sum_{i=1}^{n}V_i) \\ &= \frac{1}{n}\sum_{i=1}^{n}E(\mathcal{M}(V)+\mathcal{R}_i(V)) \\ &= \frac{1}{n}\sum_{i=1}^{n}\Sigma_A = \Sigma_A \end{aligned} \tag{5.1.1}$$

and similarly:

$$\begin{aligned} \mathrm{Var}(\bar{D}_n^{(2)}) &= \mathrm{Var}(\frac{1}{n}\sum_{i=1}^{n}(A_i-\mu)^2) = \mathrm{Var}(\frac{1}{n}\sum_{i=1}^{n}V_i) \\ &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n}(\mathcal{M}(V)+\mathcal{R}_i(V))\right) \\ &= \mathrm{Var}(\mathcal{M}(V)+\frac{1}{n}\sum_{i=1}^{n}\mathcal{R}_i(V)) \\ &= \Xi_V + \frac{1}{n}(\Phi_V-\Xi_V) \end{aligned} \tag{5.1.2}$$

and:

$$\begin{aligned} \mathrm{Cov}(\mathcal{M}(V),\bar{D}_n^{(2)}) &= \mathrm{Cov}(\mathcal{M}(V),\frac{1}{n}\sum_{i=1}^{n}(A_i-\mu_A)^2) \\ &= \mathrm{Cov}(\mathcal{M}(V),\frac{1}{n}\sum_{i=1}^{n}V_i) \\ &= \mathrm{Cov}(\mathcal{M}(V),\mathcal{M}(V)+\frac{1}{n}\sum_{i=1}^{n}\mathcal{R}_i(V))) \\ &= \Xi_V \end{aligned} \tag{5.1.3}$$

Then using equations 5.1.1, 5.1.2 and 5.1.3, the adjusted expectation for $\mathcal{M}(V)$ is given by:

$$\begin{aligned} E_n(\mathcal{M}(V)) &= E(\mathcal{M}(V)) + \mathrm{Cov}(\mathcal{M}(V),\bar{D}_n^{(2)})\mathrm{Var}(\bar{D}_n^{(2)})^{\dagger}(\bar{D}_n^{(2)}-E(\bar{D}_n^{(2)})) \\ &= \Sigma_A + \Xi_V\frac{1}{\Xi_V+\frac{1}{n}(\Phi_V-\Xi_V)}(\bar{D}^{(2)}-\Sigma_A) \\ &= \frac{\Xi_V\bar{D}_n^{(2)}+\frac{1}{n}(\Phi_V-\Xi_V)\Sigma_A}{\Xi_V+\frac{1}{n}(\Phi_V-\Xi_V)} \end{aligned} \tag{5.1.4}$$

This is a weighted average of the prior and observed estimates. As the sample size increases our weighting of the prior information decreases as we place more belief in the observed estimate.

5.2 Updating the variance of a univariate DLM

We now consider Bayes Linear variance learning for a simple univariate dynamic linear model. This model can be seen as a single component of the corrosion model in section 3.5, with fixed corrosion rate and no observation error. We have a model in time t:

$$X_t = X_{t-1} - \alpha + \epsilon_{Xt}$$

where α is a known fixed slope term. The residual ϵ_{Xt} has known mean 0 and is independent in time so our prior specification is:

$$E(\epsilon_{Xt}) = 0 \quad \text{Var}(\epsilon_{Xt}) = \Sigma_X \quad \text{Cov}(\epsilon_{Xt}, \epsilon_{Xt'}) = 0$$

Using the ideas in the previous section 5.1, we know how to update the population variance of a collection of quantities with known mean. In our DLM, the set of $\{\epsilon_{Xt}\}$ are a collection of quantities which have prior expected value 0 and prior variance Σ_X. We make the assumption that the residuals ϵ_{Xt} are second order exchangeable in time giving a representation:

$$\epsilon_{Xt} = \mathcal{M}(\epsilon_X) + \mathcal{R}_t(\epsilon_X)$$

In this case we know that $\mathcal{M}(\epsilon_X) = 0$ so the representation reduces to:

$$\epsilon_{Xt} = \mathcal{R}_t(\epsilon_X) \tag{5.2.5}$$

where $\text{Var}(\mathcal{R}_t(\epsilon_X)) = \Sigma_X$. We also assume that the squared residuals form are second order exchangeable in time with prior specification:

$$E(\mathcal{M}(V)) = \Sigma_X \qquad \text{Var}(\mathcal{M}(V)) = \Xi_V \qquad \text{Var}(\mathcal{R}_t(V)) = \Phi_V - \Xi_V,$$

giving a representation:

$$(\mathcal{R}_t(\epsilon_X))^2 = V_t = \mathcal{M}(V) + \mathcal{R}_t(V)$$

Using observations from our system we wish to update our beliefs about our "population" variance $\mathcal{M}(V)$.

We do not directly observe the residuals. However, using a simple linear combination of time steps, we can get a an expression for the residuals at every point in time:

$$\epsilon_{Xt} = X_t - X_{t-1} + \alpha.$$

We can then find:

$$\begin{aligned}\bar{D_T} &= \frac{1}{T-1}\sum_{t=1}^{T-1}(X_t - X_{t-1} + \alpha)^2 \\ &= \frac{1}{T-1}\sum_{t=1}^{T-1}\epsilon_{Xt}^2\end{aligned}$$

where T is the total number of time observed points. Then as in equations 5.1.1,5.1.2 and 5.1.3 we can find that:

$$\begin{aligned}E(\bar{D_T}) &= \Sigma_X \\ \mathrm{Var}(\bar{D_T}) &= \Xi_V + \frac{1}{T-1}(\Phi_V - \Xi_V) \\ \mathrm{Cov}(\mathcal{M}(V), D_T) &= \Xi_V\end{aligned}$$

and using the same approach as for equation 5.1.4 we find that the adjusted expectation, $E_T(\mathcal{M}(V))$, is given by:

$$E_T(\mathcal{M}(V)) = \frac{\Xi_V \bar{D_T} + \frac{1}{T-1}(\Phi_V - \Xi_V)\Sigma_X}{\Xi_V + \frac{1}{T-1}(\Phi_V - \Xi_V)}$$

5.3 Bayes linear variance learning for the linear growth DLM

We have shown how to update a population variance and residual variance in a univariate DLM. We now consider a learning about the residual variances for a multivariate linear growth DLM. This model is used in [Randell et al., 2010] to carry out variance learning and design inspection schemes. This paper won the 2010 Donald Julius Groen Prize from the safety and reliability group of the Institution of Mechanical Engineers.

5.3.1 Linear growth model

Consider inspection of a collection of components over time. A linear growth DLM, is used for the system level, X_{ct}, and system slope α_{ct}, for component, c, at time t. Observations of the system state, Y_{ct}, are made subject to measurement error of the form, ϵ_{Yct}. The model equations are:

$$\begin{aligned} \text{Observation:} \qquad & Y_{ct} = X_{ct} + \epsilon_{Yct} \\ \text{System level:} \qquad & X_{ct} = X_{c(t-1)} + \alpha_{ct} + \epsilon_{Xct} \\ \text{System slope:} \qquad & \alpha_{ct} = \alpha_{c(t-1)} + \epsilon_{\alpha ct}, \end{aligned} \tag{5.3.6}$$

where the sets of residual $\{\epsilon_{Yct}\}$, $\{\epsilon_{Xct}\}$ and $\{\epsilon_{\alpha ct}\}$ are mutually uncorrelated across groups. We make a prior specification about the residuals as follows:

$$\begin{aligned} E[\epsilon_{Yct}] &= 0 & E[\epsilon_{Xct}] &= 0 & E[\epsilon_{\alpha ct}] &= 0 \\ \text{Var}[\epsilon_{Yct}] &= \Sigma_Y & \text{Var}[\epsilon_{Xct}] &= \Sigma_{Xc} & \text{Var}[\epsilon_{\alpha ct}] &= \Sigma_{\alpha c}. \end{aligned}$$

We assume that the residuals are independent in time but correlated across components:

$$\begin{aligned} \text{Cov}[\epsilon_{Yct}, \epsilon_{Yc't}] &= \Sigma_{Ycc'} \quad c \neq c' & \text{Cov}[\epsilon_{Yct}, \epsilon_{Yc't'}] &= 0 \quad t \neq t' \quad \forall c, c' \\ \text{Cov}[\epsilon_{Xct}, \epsilon_{Xc't}] &= \Sigma_{Xcc'} \quad c \neq c' & \text{Cov}[\epsilon_{Xct}, \epsilon_{Xc't'}] &= 0 \quad t \neq t' \quad \forall c, c' \\ \text{Cov}[\epsilon_{\alpha ct}, \epsilon_{\alpha c't}] &= \Sigma_{\alpha cc'} \quad c \neq c' & \text{Cov}[\epsilon_{\alpha ct}, \epsilon_{\alpha c't'}] &= 0 \quad t \neq t' \quad \forall c, c'. \end{aligned} \tag{5.3.7}$$

The evolution of this system is driven by the residuals ϵ_{Xct} and $\epsilon_{\alpha ct}$ and their underlying (co)variance structure. Therefore learning about these variances is especially important for reliable forecasting.

5.3.2 Exchangeability in time

We make exchangeability assumptions for the residuals as in section 5.2, the major differences being that we have three residuals and we now consider multivariate quantities.

Consider the system level residual ϵ_{Xct}, we make the assumptions that for each component both the residuals and the squared residuals are second order exchangeable in time as in equation 5.2.5. This gives representation statements for each

component:

$$\epsilon_{Xct} = \mathcal{R}_t(\epsilon_{Xc})$$
$$(\mathcal{R}_t(\epsilon_{Xc}))^2 = V_{Xct} = \mathcal{M}(V_{Xc}) + \mathcal{R}_t(V_{Xc}) \tag{5.3.8}$$

where we make the prior specification:

$$E(\mathcal{M}(V_{Xc})) = \Sigma_{Xc} \quad \text{Var}(\mathcal{M}(V_{Xc})) = \Xi_{V_{Xc}} \quad \text{Var}(\mathcal{R}_t(V_{Xc})) = \Phi_{V_{Xc}} - \Xi_{V_{Xc}},$$

Then to update the variance of the system level residual of particular components we learn about $\mathcal{M}(V_{Xc})$.

We can make similar exchangeability assumptions about the other residuals, giving representations for $\epsilon_{\alpha ct}$:

$$\epsilon_{\alpha ct} = \mathcal{R}_t(\epsilon_{\alpha c})$$
$$(\mathcal{R}_t(\epsilon_{\alpha c}))^2 = V_{\alpha ct} = \mathcal{M}(V_{\alpha c}) + \mathcal{R}_t(V_{\alpha c}) \tag{5.3.9}$$

and for ϵ_{Yct}:

$$\epsilon_{Yct} = \mathcal{R}_t(\epsilon_{Yc})$$
$$(\mathcal{R}_t(\epsilon_{Yc}))^2 = V_{Yct} = \mathcal{M}(V_{Yc}) + \mathcal{R}_t(V_{Yc}) \tag{5.3.10}$$

5.3.3 Squared linear combinations of observations

Within the DLM it is the residuals and their variances which drive the evolution of the system. However we don't directly observe these residuals. We consider squared linear combinations of observations which eliminate the slope and level terms. This results in expressions which only involve residual terms. These are then informative for variance learning.

Assuming full inspections observations with equally spaced in time; for component, c let:

$$Y_{ct}^{(i)} = Y_{ct} - Y_{c(t-i)}$$

be the ith time step differences of observations. Using equation 5.3.6 and taking one

step time difference, $Y_{ct}^{(1)}$, we find:

$$\begin{aligned} Y_{ct}^{(1)} = Y_{ct} - Y_{c(t-1)} &= X_{ct} - X_{c(t-1)} + \epsilon_{Yct} - \epsilon_{Yc(t-1)} \\ &= \alpha_{ct} + \epsilon_{Xct} + \left(\epsilon_{Yct} - \epsilon_{Yc(t-1)}\right) \\ &= \alpha_{c(t-1)} + \epsilon_{\alpha ct} + \epsilon_{Xct} + \left(\epsilon_{Yct} - \epsilon_{Yc(t-1)}\right) \end{aligned} \tag{5.3.11}$$

and similarly we can find the two step differences, $Y_{ct}^{(2)}$:

$$\begin{aligned} Y_{ct}^{(2)} = Y_{ct} - Y_{c(t-2)} &= X_{ct} - X_{c(t-2)} + \epsilon_{Yct} - \epsilon_{Yc(t-2)} \\ &= \alpha_{ct} + X_{c(t-1)} - X_{c(t-2)} + \epsilon_{Xct} + \epsilon_{Yct} - \epsilon_{Yc(t-2)} \\ &= \alpha_{ct} + \alpha_{c(t-1)} + \epsilon_{Xct} + \epsilon_{Xc(t-1)} + \epsilon_{Yct} - \epsilon_{Yc(t-2)} \\ &= 2\alpha_{c(t-1)} + \epsilon_{\alpha ct} + \epsilon_{Xct} + \epsilon_{Xc(t-1)} + \epsilon_{Yct} - \epsilon_{Yc(t-2)}. \end{aligned} \tag{5.3.12}$$

Using equations 5.3.11 and 5.3.12 gives the linear combination of differences:

$$Y_{ct}^{(2)} - 2Y_{ct}^{(1)} = -\epsilon_{\alpha ct} + \epsilon_{Xct} - \epsilon_{Xc(t-1)} + 2\epsilon_{Yc(t-1)} - \epsilon_{Yct} - \epsilon_{Yc(t-2)}. \tag{5.3.13}$$

This gives an expression involving only the residuals, the square of which is informative for variance learning. The particular linear combination, is obviously, not unique, just one such combination which removes system level effects.

5.3.4 Updating system variances for individual components

We now consider updating beliefs about system level residual variance, $\mathrm{Var}(\epsilon_{Xct})$, using exchangeability assumptions and observations on that component. Let D_{ct} be the linear combination:

$$\begin{aligned} D_{ct} &= (Y_{ct}^{(2)} - 2Y_{ct}^{(1)})^2 \\ &= \left(\left(Y_{ct} - Y_{c(t-2)}\right) - 2\left(Y_{ct} - Y_{c(t-1)}\right)\right)^2 \end{aligned} \tag{5.3.14}$$

If $t = 1$ is the start time, then for component c let D_c:

$$D_c = \begin{pmatrix} D_{c3} \\ \vdots \\ D_{cT} \end{pmatrix} \tag{5.3.15}$$

Theorem 5.3.1 The adjusted expectation of $E_{D_c}(\mathcal{M}(V_{Xc}))$ is:

$$E_{D_c}(\mathcal{M}(V_{Xc})) = \Sigma_{Xc} + 2\Xi_{V_{Xc}}(\mathrm{Var}(D_c))^{-1}(D_c - 1_T(\Sigma_{\alpha c} + 2\Sigma_{Xc} + 6\Sigma_Y))$$

where $\mathrm{Var}(D_c)$ is a matrix composed of elements from equations 5.3.18, 5.3.19, 5.3.20 and 5.3.21 and 1_T is the $T \times 1$ vector of ones.

Proof:

Using equation 5.3.13 we find the expectation of D_{ct} for all, c and t:

$$\begin{aligned} E[D_{ct}] &= E\left[(Y_{ct}^{(2)} - 2Y_{ct}^{(1)})^2\right] \\ &= E\left[(-\epsilon_{\alpha ct} + \epsilon_{Xct} - \epsilon_{Xc(t-1)} + 2\epsilon_{Yc(t-1)} - \epsilon_{Yct} - \epsilon_{Yc(t-2)})^2\right] \\ &= E\left[(-\epsilon_{\alpha ct})^2\right] + E\left[(\epsilon_{Xct} - \epsilon_{Xc(t-1)})^2\right] \\ &+ E\left[\left(2\epsilon_{Yc(t-1)} - \epsilon_{Yct} - \epsilon_{Yc(t-2)}\right)^2\right] \\ &= \Sigma_{\alpha c} + 2\Sigma_{Xc} + 6\Sigma_Y; \end{aligned} \tag{5.3.16}$$

as the terms involving cross products are all zero. Then using equations, 5.3.8, 5.3.9 and 5.3.10 we find:

$$\begin{aligned} \mathrm{Cov}(\mathcal{M}(V_{Xc}), D_{ct}) &= \mathrm{Cov}\left(\mathcal{M}(V_{Xc}), \left(-\epsilon_{\alpha ct} + \epsilon_{Xct} - \epsilon_{Xc(t-1)}\right.\right. \\ &\qquad \left.\left. +2\epsilon_{Yc(t-1)} - \epsilon_{Yct} - \epsilon_{Yc(t-2)}\right)^2\right) \\ &= \mathrm{Cov}\left(\mathcal{M}(V_{Xc}), (\mathcal{M}(V_{\alpha c}) + \mathcal{R}_t(V_{\alpha c}) + 2\mathcal{M}(V_{Xc}) + \mathcal{R}_t(V_{Xc})\right. \\ &\left. +\mathcal{R}_{t-1}(V_{Xc}) + 6\mathcal{M}(V_Y) + 4\mathcal{R}_{t-1}(V_Y) + \mathcal{R}_{t-2}(V_Y) + \mathcal{R}_t(V_Y))\right) \\ &= 2\mathrm{Cov}\left(\mathcal{M}(V_{Xc}), \mathcal{M}(V_{Xc})\right) \\ &= 2\Xi_{V_{Xc}} \end{aligned} \tag{5.3.17}$$

We can also find $\mathrm{Var}[D_{ct}]$ using equations 5.3.13, 5.3.8, 5.3.9 and 5.3.10:

$$\begin{aligned} \mathrm{Var}[D_{ct}] &= \mathrm{Var}\left(-\epsilon_{\alpha ct} + \epsilon_{Xct} - \epsilon_{Xc(t-1)} + 2\epsilon_{Yc(t-1)} - \epsilon_{Yct} - \epsilon_{Yc(t-2)}\right)^2 \\ &= \mathrm{Var}\left(\mathcal{M}(V_{\alpha c}) + \mathcal{R}_t(V_{\alpha c}) + 2\mathcal{M}(V_{Xc}) + \mathcal{R}_t(V_{Xc})\right. \\ &\left. +\mathcal{R}_{t-1}(V_{Xc}) + 6\mathcal{M}(V_Y) + 4\mathcal{R}_{t-1}(V_Y) + \mathcal{R}_{t-2}(V_Y) + \mathcal{R}_t(V_Y)\right) \\ &= \Phi_{V_{\alpha c}} + 2\Xi_{V_{Xc}} + 2\Phi_{V_{Xc}} + 18\Xi_{V_Y} + 18\Phi_{V_Y} \end{aligned} \tag{5.3.18}$$

Similarly the covariance between squared linear combinations at different times can be found:

$$\text{Cov}\left[D_{ct}, D_{c(t-1)}\right] = \Xi_{V_{\alpha c}} + 3\Xi_{V_{Xc}} + \Phi_{V_{Xc}} + 28\Xi_{V_Y} + 8\Phi_{V_Y} \tag{5.3.19}$$

$$\text{Cov}\left[D_{ct}, D_{c(t-2)}\right] = \Xi_{V_{\alpha c}} + 4\Xi_{V_{Xc}} + 35\Xi_{V_Y} + \Phi_{V_Y} \tag{5.3.20}$$

$$\text{Cov}\left[D_{ct}, D_{c(t-k)}\right] = \Xi_{V_{\alpha c}} + 4\Xi_{V_{Xc}} + 36\Xi_{V_Y} \qquad \text{where } k \geq 3. \tag{5.3.21}$$

Using equations 5.3.16 and 5.3.17, the adjusted expectation is given by:

$$\begin{aligned} E_{D_c}(\mathcal{M}(V_{Xc})) &= E(\mathcal{M}(V_{Xc})) + \text{Cov}(\mathcal{M}(V_{Xc}), D_c)(\text{Var}(D_c))^{-1}(D_c - E(D_c)) \\ &= \Sigma_{Xc} + 2\Xi_{V_{Xc}}(\text{Var}(D_c))^{-1}(D_c - 1_T(\Sigma_{\alpha c} + 2\Sigma_{Xc} + 6\Sigma_Y)) \end{aligned}$$

where $(\text{Var}(D_c))$ is a matrix composed of elements from equations 5.3.18, 5.3.19, 5.3.20 and 5.3.21. □

Similar expressions can easily be found for adjusting beliefs regarding the other system variances $\Sigma_{\alpha c}$ and Σ_Y using:

$$\begin{aligned} \text{Cov}(\mathcal{M}(V_Y), D_{ct}) &= \text{Cov}(\mathcal{M}(V_Y), D_{ct}) \\ &= \text{Cov}\left(\mathcal{M}(V_Y), (\mathcal{M}(V_{\alpha c}) + \mathcal{R}_t(V_{\alpha c}) + 2\mathcal{M}(V_{Xc}) + \mathcal{R}_t(V_{Xc})\right. \\ &\left. + \mathcal{R}_{t-1}(V_{Xc}) + 6\mathcal{M}(V_Y) + 4\mathcal{R}_{t-1}(V_Y) + \mathcal{R}_{t-2}(V_Y) + \mathcal{R}_t(V_Y))\right) \\ &= 6\Xi_{V_Y}, \end{aligned}$$

and:

$$\begin{aligned} \text{Cov}(\mathcal{M}(V_{\alpha c}), D_{ct}) &= \text{Cov}(\mathcal{M}(V_{\alpha c}), D_{ct}) \\ &= \text{Cov}\left(\mathcal{M}(V_{\alpha c}), (\mathcal{M}(V_{\alpha c}) + \mathcal{R}_t(V_{\alpha c}) + 2\mathcal{M}(V_{Xc}) + \mathcal{R}_t(V_{Xc})\right. \\ &\left. + \mathcal{R}_{t-1}(V_{Xc}) + 6\mathcal{M}(V_Y) + 4\mathcal{R}_{t-1}(V_Y) + \mathcal{R}_{t-2}(V_Y) + \mathcal{R}_t(V_Y))\right) \\ &= \Xi_{V_{\alpha c}}, \end{aligned}$$

therefore the adjusted expectation,$E_{D_c}(\mathcal{M}(V_Y))$ and $E_{D_c}(\mathcal{M}(V_{\alpha c}))$ are given by:

$$\begin{aligned} E_{D_c}(\mathcal{M}(V_{\alpha c})) &= E(\mathcal{M}(V_{\alpha c})) + \text{Cov}(\mathcal{M}(V_{\alpha c}), D_c)(\text{Var}(D_c))^{-1}(D_c - E(D_c)) \\ &= \Sigma_{\alpha c} + \Xi_{V_{\alpha c}}(\text{Var}(D_c))^{-1}(D_c - 1_T(\Sigma_{\alpha c} + 2\Sigma_{Xc} + 6\Sigma_Y)) \end{aligned}$$

and:

$$E_{D_c}(\mathcal{M}(V_Y)) = E(\mathcal{M}(V_Y)) + \text{Cov}(\mathcal{M}(V_Y), D_c)(\text{Var}(D_c))^{-1}(D_c - E(D_c))$$
$$= \Sigma_Y + 6\Xi_{V_Y}(\text{Var}(D_c))^{-1}(D_c - 1_T(\Sigma_{\alpha c} + 2\Sigma_{Xc} + 6\Sigma_Y))$$

When adjusting variances the covariances are adjusted to maintain the same underlying correlation structure. In chapter 9 we look at approaches to learning about the full covariance structure.

5.4 Exchangeability of variances across components

If we are willing to express prior beliefs about certain fourth order quantities we now have a method, in section 5.3.4, for learning about the variances within the model. Given long time series of complete regularly spaced observations this would most likely be sufficient to give good estimates for the standard error variances. In our case, however we have large numbers of components each with relatively short time series. We want to be able to share information across components within the system. To do this we need expressions for beliefs about the relationship between variances within the model. This is achieved by assuming exchangeability of the variances in the model across components.

We assume second order exchangeability of mean system level evolution variance, $\mathcal{M}(V_{Xc})$, over components. This leads to representation statements for the variance of every component, $c = 1, 2, ...C$:

$$\mathcal{M}(V_{Xc}) = W_{Xc} = \mathcal{M}(W_X) + \mathcal{R}_c(W_X) \tag{5.4.22}$$

where:

$$E(\mathcal{M}(W_X)) = \Sigma_{W_X} \quad \text{Var}(W_{Xc}) = \Phi_{W_X} \quad \text{Cov}(W_X, W_{Xc'}) = \Xi_{W_X}, \quad c \neq c'$$

The adjusted expectation $E_D(\mathcal{M}(W_X))$ gives an updated estimate of, $\mathcal{M}(W_X)$ and so tells us something about $\mathcal{M}(V_{Xc})$ for all components c.

Using the second order exchangeability assumptions we can use all the data for all components so we write the $C(T-3)\times 1$, vector D:

$$D = \begin{pmatrix} D_1 \\ D_2 \\ \vdots \\ D_c \vdots \\ D_C \end{pmatrix}$$

where D_c is as given in equation 5.3.15. Then we adjust beliefs about the mean variance, $\mathcal{M}(W_X)$ which tells us about all the variances in the model. Using equations 5.3.16, 5.3.17 and 5.4.22:

$$\begin{aligned} E_D(\mathcal{M}(W_X)) &= E(\mathcal{M}(W_X)) + \mathrm{Cov}(\mathcal{M}(W_X), D)(\mathrm{Var}(D))^{-1}(D - E(D)) \\ &= \Sigma_{W_X} + 2(\Xi_{W_X}, \ldots, \Xi_{W_X})(\mathrm{Var}(D))^{-1}(D - 1_C(\Sigma_{\alpha c} + 2\Sigma_{Xc} + 6\Sigma_Y)) \end{aligned}$$

where $\mathrm{Var}(D)$ can be found using equations 5.3.18, 5.3.19, 5.3.20 and 5.3.21, where 1_C is the $C\times 1$ vector of ones The adjusted variance is then:

$$\begin{aligned} \mathrm{Var}_D(\mathcal{M}(W_X)) &= \mathrm{Var}(\mathcal{M}(W_X)) - \mathrm{Cov}(\mathcal{M}(W_X), D)(\mathrm{Var}(D))^{-1}\mathrm{Cov}(D, \mathcal{M}(W_X)) \\ &= \Xi_{W_X} - 4(\Xi_{W_X}, \ldots, \Xi_{W_X})(\mathrm{Var}(D))^{-1}(\Xi_{W_X}, \ldots, \Xi_{W_X})' \end{aligned}$$

The use of a second order exchangeability assumption across variances allows us to have a simple way of adjusting beliefs all at once. Whether this is a sensible assumption to make is another question. We need to express relationships between large numbers of components. In certain cases, it may be preferable to express exchangeability assumptions across subsets of the variances which we thought a priori were behaving similarly. We could then adjust beliefs using those subsets. The other variances can be updated using a similar approach.

5.5 Irregularly Spaced Observations

We have so far shown how to update variances in the case of full inspections. In reality we have irregularly spaced partial inspections so we need to be able to handle this type of data.

5.5.1 Matrix form of the linear growth DLM

In the case of irregular and partial inspections we can find similar types of linear combinations to those discussed in section 5.3. Consider the model in equation 5.3.6. The DLM can be written in matrix form and generalised to give an expression for the evolution of the DLM with general time steps.

Lemma 5.5.1 The linear growth DLM can be written in the form:

$$Y_{ct} = \begin{pmatrix} 1 & 0 \end{pmatrix} \Theta_{ct} + \epsilon_{Yct}$$
$$\Theta_{ct} = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \Theta_{c(t-k)} + \xi_{c(t,t-k)} \quad \forall k$$

where:

$$\Theta_{ct} = \begin{pmatrix} X \\ \alpha \end{pmatrix}_{ct} \qquad \text{and} \qquad \xi_{c(t,t-k)} = \sum_{i=0}^{k-1} \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{ct-i}$$

Proof: Let:

$$\Theta_{ct} = \begin{pmatrix} X \\ \alpha \end{pmatrix}_{ct}$$

then from equation 5.3.6:

$$\begin{aligned} Y_{ct} &= X_{ct} + \epsilon_{Yct} \\ &= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ \alpha \end{pmatrix}_{ct} + \epsilon_{Yct} \\ &= \begin{pmatrix} 1 & 0 \end{pmatrix} \Theta_{ct} + \epsilon_{Yct} \end{aligned}$$

and from equation 5.3.6:

$$X_{ct} = X_{c(t-1)} + \alpha_{c(t-1)} + \epsilon_{\alpha ct} + \epsilon_{Xct}$$
$$\alpha_{ct} = \alpha_{c(t-1)} + \epsilon_{\alpha ct}$$

in matrix form this can be written as:

$$\begin{aligned}
\Theta_{ct} &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \Theta_{c(t-1)} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{ct} \\
&= \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \Theta_{c(t-2)} + \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{c(t-1)} \\
&+ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{ct} \\
&= \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \Theta_{c(t-k)} + \sum_{i=0}^{k-1} \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{c(t-i)} \\
&= \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \Theta_{c(t-k)} + \xi_{c(t,t-k)}
\end{aligned}$$

□

This general form of the DLM has a more complicated form of the residuals. We can still find its mean and variance in terms of the original specification:

$$\begin{aligned}
E(\xi_{c(t,t-k)}) &= E\left[\sum_{i=0}^{k-1} \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{c(t-i)} \right] \\
&= \sum_{i=0}^{k-1} \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} E\left[\begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{c(t-i)} \right] \\
&= \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\end{aligned}$$

and its variance:

$$\begin{aligned}
\mathrm{Var}[\xi_{c(t,t-k)}] &= \sum_{i=0}^{k-1} \mathrm{Var}\left[\begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{ct-i}\right] \\
&= \sum_{i=0}^{k-1} \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \mathrm{Var}\begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{ct-i} \begin{pmatrix} 1 & 0 \\ i & 1 \end{pmatrix} \\
&= \sum_{i=0}^{k-1} \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Sigma_{Xc} + \Sigma_{\alpha c} & \Sigma_{\alpha c} \\ \Sigma_{\alpha c} & \Sigma_{\alpha c} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ i & 1 \end{pmatrix} \\
&= \sum_{i=0}^{k-1} \begin{pmatrix} \Sigma_{Xc} + (1+i)^2\Sigma_{\alpha c} & (1+i)\Sigma_{\alpha c} \\ (1+i)\Sigma_{\alpha c} & \Sigma_{\alpha c} \end{pmatrix} \\
&= \begin{pmatrix} k\Sigma_{Xc} + \frac{1}{6}k(k+1)(2k+1)\Sigma_{\alpha c} & \frac{1}{2}k(k+1)\Sigma_{\alpha c} \\ \frac{1}{2}k(k+1)\Sigma_{\alpha c} & k\Sigma_{\alpha c} \end{pmatrix} \qquad (5.5.23)
\end{aligned}$$

5.5.2 Squared linear combinations of observations

We use the general form of the DLM to find linear combinations of observations which eliminate the slope and level terms. This leaves expressions only involving the residual structure as in section 5.3.3.

Consider a set of irregularly spaced observations for particular component, c:

$$\{Y_{ct_1}, Y_{ct_2}, \ldots, Y_{ct_{T_c}}\}$$

observed at time:

$$\{t_1, t_2, \ldots, t_{T_c}\}.$$

We consider sets of 3 observations at times Y_{ct}, $Y_{c(t_i-k_i)}$ and $Y_{c(t_i-l_i)}$ where:

$$k_i = t_i - t_{i-1} \qquad \text{and} \qquad l_i = t_i - t_{i-2}$$

here $k_i < l_i$, these represent the three most recent previous observations of the component. Then as in section 5.3.3 we look at differences of observations. Using

lemma 5.5.1 we find:

$$\begin{aligned}
Y_c^{(k_i)} &= Y_{ct_i} - Y_{c(t_i-k_i)} \\
&= \begin{pmatrix} 1 & 0 \end{pmatrix} \Theta_{ct_i} - \begin{pmatrix} 1 & 0 \end{pmatrix} \Theta_{c(t_i-k_i)} + \epsilon_{Yct_i} - \epsilon_{Yc(t_i-k_i)} \\
&= \begin{pmatrix} 1 & k_i \end{pmatrix} \Theta_{c(t_i-k_i)} - \begin{pmatrix} 1 & 0 \end{pmatrix} \Theta_{c(t_i-k_i)} \\
&\quad + \begin{pmatrix} 1 & 0 \end{pmatrix} \xi_{c(t_i,t_i-k_i)} + \epsilon_{Yct_i} - \epsilon_{Yc(t_i-k_i)} \\
&= \begin{pmatrix} 0 & k_i \end{pmatrix} \Theta_{c(t_i-k_i)} + \begin{pmatrix} 1 & 0 \end{pmatrix} \xi_{c(t_i,t_i-k_i)} + \epsilon_{Yct_i} - \epsilon_{Yc(t_i-k_i)} \\
&= \begin{pmatrix} 0 & k_i \end{pmatrix} \Theta_{c(t_i-l_i)} + \begin{pmatrix} 0 & k_i \end{pmatrix} \xi_{c(t_i-k_i,t_i-l_i)} \\
&\quad + \begin{pmatrix} 1 & 0 \end{pmatrix} \xi_{c(t_i,t_i-k_i)} + \epsilon_{Yct_i} - \epsilon_{Yc(t_i-k_i)}
\end{aligned} \tag{5.5.24}$$

and:

$$\begin{aligned}
Y_c^{(l_i)} &= Y_{ct_i} - Y_{c(t_i-l_i)} \\
&= \begin{pmatrix} 1 & 0 \end{pmatrix} \Theta_{ct_i} - \begin{pmatrix} 1 & 0 \end{pmatrix} \Theta_{c(t_i-l_i)} + \epsilon_{Yct_i} - \epsilon_{Yc(t_i-l_i)} \\
&= \begin{pmatrix} 1 & k_i \end{pmatrix} \Theta_{c(t_i-k_i)} - \begin{pmatrix} 1 & 0 \end{pmatrix} \Theta_{c(t_i-l_i)} \begin{pmatrix} 1 & 0 \end{pmatrix} \xi_{c(t_i,t_i-k_i)} \\
&\quad + \epsilon_{Yct_i} - \epsilon_{Yc(t_i-l_i)} \\
&= \begin{pmatrix} 0 & l_i \end{pmatrix} \Theta_{c(t_i-l_i)} + \begin{pmatrix} 1 & k_i \end{pmatrix} \xi_{c(t_i-k_i,t_i-l_i)} + \begin{pmatrix} 1 & 0 \end{pmatrix} \xi_{c(t_i,t_i-k_i)} \\
&\quad + \epsilon_{Yct_i} - \epsilon_{Yc(t_i-l_i)}
\end{aligned} \tag{5.5.25}$$

We can now find a linear combination of the observations which eliminates the Θ_{ct} terms using equations 5.5.24 and 5.5.25. This will leave an expression involving only the residuals:

$$\begin{aligned}
k_i Y_c^{(l_i)} - l_i Y_c^{(k_i)} &= k_i \Big[\begin{pmatrix} 0 & l_i \end{pmatrix} \Theta_{c(t_i-l_i)} + \begin{pmatrix} 1 & k_i \end{pmatrix} \xi_{c(t_i-k_i,t_i-l_i)} \\
&\quad + \begin{pmatrix} 1 & 0 \end{pmatrix} \xi_{c(t_i,t_i-k_i)} + \left(\epsilon_{Yct_i} - \epsilon_{Yc(t_i-l_i)}\right) \Big] \\
&\quad - l_i \Big[\begin{pmatrix} 0 & k_i \end{pmatrix} \Theta_{c(t_i-l_i)} + \begin{pmatrix} 0 & k_i \end{pmatrix} \xi_{c(t_i-k_i,t_i-l_i)} \\
&\quad + \begin{pmatrix} 1 & 0 \end{pmatrix} \xi_{c(t_i,t_i-k_i)} + \left(\epsilon_{Yct_i} - \epsilon_{Yc(t_i-l_i)}\right) \Big] \\
&= \begin{pmatrix} (k_i - l_i) & 0 \end{pmatrix} \xi_{c(t_i,t_i-k_i)} + \begin{pmatrix} k_i & k_i(k_i - l_i) \end{pmatrix} \xi_{c(t_i-k_i,t_i-l_i)} \\
&\quad + k_i \left(\epsilon_{Yct_i} - \epsilon_{Yc(t_i-l_i)}\right) - l_i \left(\epsilon_{Yct_i} - \epsilon_{Yc(t_i-k_i)}\right)
\end{aligned} \tag{5.5.26}$$

Then if we look at the squared expectation of 5.5.26 using equation 5.5.23:

$$\begin{aligned}E[(k_iY_c^{(l_i)} - l_iY_c^{(k_i)})^2] =&(k_i - l_i)^2\left[k_i\Sigma_X + \frac{\Sigma_\alpha k_i(k_i+1)(2k_i+1)}{6}\right]\\ &+ k_i^2\left[(l_i-k_i)\Sigma_X + \frac{\Sigma_\alpha(l_i-k_i)(l_i-k_i+1)(2(l_i-k_i)+1)}{6}\right]\\ &- k_i^2(k_i-l_i)^2(l_i-k_i+1)\Sigma_\alpha\\ &+ k_i^2(k_i-l_i)^2(l_i-k_i)\Sigma_\alpha\\ &+ \left(2l_i^2 - 2k_il_i + 2k_i^2\right)\Sigma_Y\\ =&\frac{k_il_i(k_i-l_i)(2k_i^2-2k_il_i-1)}{6}\sigma_\alpha + k_il_i(l_i-k_i)\Sigma_X\\ &+ 2\left(l_i^2 - k_il_i + k_i^2\right)\Sigma_Y \qquad (5.5.27)\end{aligned}$$

we now have an expression involving the variances which is informative for variance learning. This is one particular choice of linear combination and is in no way unique. The use of different linear combinations or linear combinations using more observations may give more efficient estimates.

5.5.3 Updating system variances

To update our beliefs about the variance we compute $E_D(\mathcal{M}(W_X))$ (using exchangeability assumptions as in section 5.4), where:

$$D_{ct} = (k_iY_c^{(l_i)} - l_iY_c^{(k_i)})^2, \qquad (5.5.28)$$

and:

$$D = \begin{pmatrix} D_{11} \\ \vdots \\ D_{CT} \end{pmatrix}.$$

Then the adjusted expectation $E_D(\mathcal{M}(W_X))$ is given by:

$$E_D(\mathcal{M}(W_X)) \;=\; E(\mathcal{M}(W_X)) + \mathrm{Cov}(\mathcal{M}(W_X), D)(\mathrm{Var}(D))^{-1}(D - E(D))$$

where from equation 5.5.27:

$$E(D_{ct}) = \frac{k_il_i(k_i-l_i)(2k_i^2-2k_il_i-1)}{6}\Sigma_\alpha + k_il_i(l_i-k_i)\Sigma_X + 2\left(l_i^2 - k_il_i + k_i^2\right)\Sigma_Y,$$

also from equation 5.4.22:

$$E(\mathcal{M}(W_X)) = \Sigma_{W_X}$$

and:

$$\mathrm{Cov}[\mathcal{M}(W_X), D] = (\mathrm{Cov}[\mathcal{M}(W_X), D_{11}], \mathrm{Cov}[\mathcal{M}(W_X), D_{12}], .., \mathrm{Cov}[\mathcal{M}(W_X), D_{CT}])$$

where for $3 \leq i \leq T_c$:

$$\begin{aligned}
\mathrm{Cov}[\mathcal{M}(W_X), D_{ct}] &= \mathrm{Cov}\left[\mathcal{M}(W_X), (k_i Y_c^{(l_i)} - l_i Y_c^{(k_i)})^2\right] \\
&= \mathrm{Cov}\Bigg[\mathcal{M}(W_X), \Bigg(\begin{pmatrix} (k_i - l_i) & 0 \end{pmatrix} \xi_{c(t_i, t_i - k_i)} \\
&\quad + \begin{pmatrix} k_i & k_i(k_i - l_i) \end{pmatrix} \xi_{c(t_i - k_i, t_i - l_i)} + \\
&\quad + k_i \left(\epsilon_{Yct_i} - \epsilon_{Yc(t_i - l_i)}\right) - l_i \left(\epsilon_{Yct_i} - \epsilon_{Yc(t_i - k_i)}\right) \Bigg)^2 \Bigg] \\
&= k_i l_i (l_i - k_i) \Xi_{W_X}
\end{aligned}$$

calculation of $\mathrm{Cov}[\mathcal{M}(W_X), D_{ct}]$ is given in appendix B.1. For i outside this range we cannot form a triplet as we don't have enough observations.

This means:

$$\mathrm{Cov}(\mathcal{M}(W_X), D) = \left(k_3 l_3 (l_3 - k_3), \ldots k_T l_T (l_T - k_T)\right) \Xi_{W_X}$$

So the adjusted expectation, $E_D(\mathcal{M}(W_X))$, is given as:

$$\begin{aligned}
E_D(\mathcal{M}(W_X)) &= E(\mathcal{M}(W_X)) + \mathrm{Cov}(\mathcal{M}(W_X), D)(\mathrm{Var}(D))^{-1}(D - E(D)) \\
&= \Sigma_{W_X} + \left(k_3 l_3 (l_3 - k_3), \ldots k_T l_T (l_T - k_T)\right) \Xi_{W_X} \\
&\quad \times (\mathrm{Var}(D))^{-1}(D - E(D))
\end{aligned}$$

and the adjusted variance, $\mathrm{Var}_D(\mathcal{M}(W_X))$, is:

$$\begin{aligned}
\mathrm{Var}_D(\mathcal{M}(W_X)) &= \mathrm{Var}(\mathcal{M}(W_X)) - \mathrm{Cov}(\mathcal{M}(W_X), D)(\mathrm{Var}(D))^{-1}\mathrm{Cov}(D, \mathcal{M}(W_X)) \\
&= \Xi_{W_X} - \left(k_3 l_3 (l_3 - k_3), \ldots k_T l_T (l_T - k_T)\right) \Xi_{W_X} \\
&\quad * (\mathrm{Var}(D))^{-1} \Xi_{W_X} \left(k_3 l_3 (l_3 - k_3), \ldots k_T l_T (l_T - k_T)\right)^T
\end{aligned}$$

where $\mathrm{Var}(D)$ is evaluated as in section 5.3.3, but with irregular time steps. In practise $\mathrm{Var}(D)$ become increasing difficult to calculate in closed form so alternatively

we can use simulation to estimate its effects, as discussed in section 4.6. We can generate a large number of realisations of the model in time. For each realisation we can find the data vector, D_{ct}, as in equation 5.5.28. We can then compute the variance matrix over the set of realisations to get a simulated estimate for $\mathrm{Var}(D)$.

5.6 Summary

We have shown in section 5.3, how to update variances for a linear growth DLM using second order exchangeability assumptions in time and across components. Squared linear combinations of observations (section 5.3.3) can be used to remove system effects and isolate the residual structure to allow variance learning.

The approach of taking differences of observations was extended in section 5.5, to arbitrary and unevenly spaced inspections, using a generalised form of the DLM. Linear combinations which involve just 3 observations have been presented, this approach could easily be extended to linear combinations involving more data.

5.7 Examples

We now consider 2 examples. Firstly, an example using synthetic data, learning about the variances for a single component in the case of full inspections, as described in section 5.3. Secondly, we continue the offshore platform example (previously discussed in sections 1.4.1, 3.5 and 4.9), and consider variance learning in the case of short time series with irregular partial inspections. In these examples we consider variance learning for the linear growth DLM.

5.7.1 Updating the variance of a single component: full inspections

We first consider an example using synthetic data simulated from a linear growth DLM as described in equation 5.3.6, with one component and complete inspections. We consider an example with 100 observations. In this case, from equation 5.3.14

we can form a data vector of 98 differences,

$$D = \begin{pmatrix} (Y_3^{(2)} - 2Y_3^{(1)})^2 \\ \vdots \\ (Y_{100}^{(2)} - 2Y_{100}^{(1)})^2 \end{pmatrix}$$

Synthetic Data

We simulate 100 time points of synthetic "real" data (simulated using the prior mean specification (corrosion rates and wall thicknesses) as used in for the offshore platform example 4.9) from a linear growth DLM using variance specification;

$$\Sigma_Y = 1 \qquad \Sigma_X = 4 \qquad \Sigma_\alpha = 0.1$$

The advantage of using synthetic data is that we can analyse how well a method is performing. Figure 5.1 shows the data vector of linear combinations, D for each of the 98 together with the prior discrepancy. There are several values with very large discrepancies even in this case using synthetic data. Each point represents a particular linear combination D_t and the red line shows $E(D)$. There are several linear combinations for which D_t is large and have large discrepancies. There are several squared linear combinations for which D_t is large and consequently have large discrepancies.

Prior specification

Our prior variance specification is given as:

$$\Sigma_Y = 1 \qquad \Sigma_X = 1 \qquad \Sigma_\alpha = 0.1$$

The variance of the corrosion rate is expected to be smaller than that of the wall thickness. The prior measurement error variance, Σ_Y, and corrosion rate variance, Σ_α are correct. However the prior wall thickness variance has been underestimated. In this example we will investigate how well we can recover the "real" system variances.

In order to carry out Bayes linear variance learning we also need to specify 4th order quantities. We assume the measurement error is know and not random so that

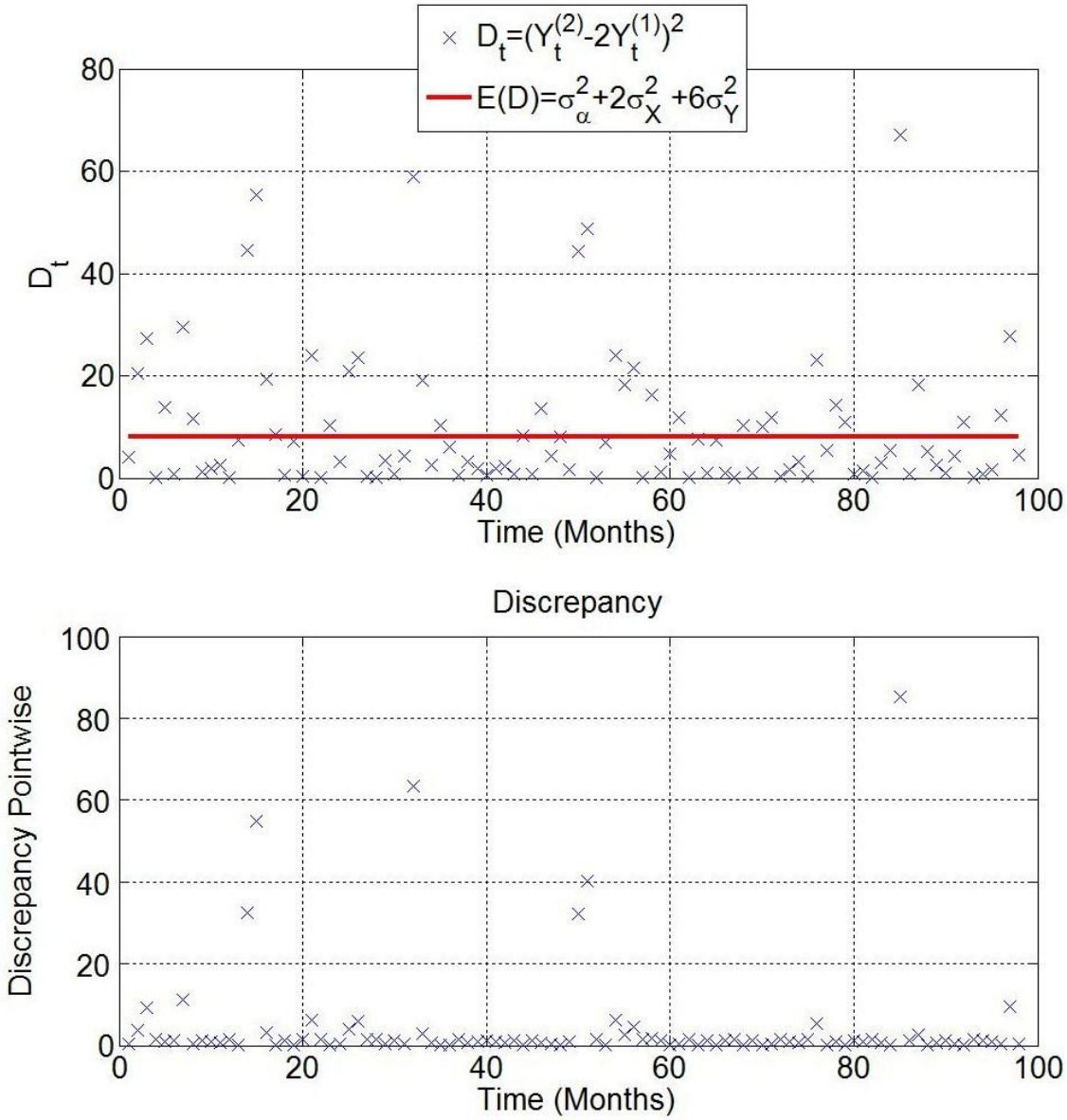

Figure 5.1: Vector of differences, D, and $E(D)$ and point-wise discrepancies,Dr(D), for 100 observations of the synthetic "real" data. Each point represents a particular linear combination D_t and the red line shows $E(D)$. There are several linear combinations for which D_t is large and consequently have large discrepancies.

$\Phi_{W_Y} = \Xi_{W_Y} = 0.$

$$\Phi_{W_Y} = 0 \qquad \Phi_{W_X} = 1 \qquad \Phi_{W_\alpha} = 0.1$$
$$\Xi_{W_Y} = 0 \qquad \Xi_{W_X} = 0.5 \qquad \Xi_{W_\alpha} = 0.05$$

Bayes linear updating

We compute the adjusted expectation, $E_D(\mathcal{M}(W_X))$, as in theorem 5.3.1.

We generate 50 sets of synthetic real data and compute the adjusted expectation, $E_D(\mathcal{M}(W_X))$, for each. Figure 5.2 shows the distribution of, $E_D(\mathcal{M}(W_X))$, in this case. We see that it is centred on the prior side of the true value for the wall thickness variance. Figure 5.3 shows the adjusted beliefs with increasing T. Prediction of $E_D(\mathcal{M}(W_X))$, is centered at the correct value and the uncertainty around our prediction decreases as we see more data. For very small numbers of observations we can see it is possible to get predictions of negative variances which is undesirable. The Bayes linear adjustment does not preclude this possibility.

5.7.2 Updating the variances for offshore platform application

We now consider an example using the offshore platform inspection application described in section 1.4.1. We generate synthetic data with the inspection design as given in figure 1.4. The advantages of using synthetic data are discussed in section 1.4.2. The data is generated using the linear growth model 5.5 with known variances. As such we can assess how well our method performs.

Synthetic data

We have 174 observations for 64 components over 83 time points from 4 corrosion circuits, unlike in the example in section 5.7.1, where we considered full inspections. In this case we have short time series of data per component with incomplete and irregularly spaced inspections. Figure 5.4 shows the synthetic data with the observations shown in red.

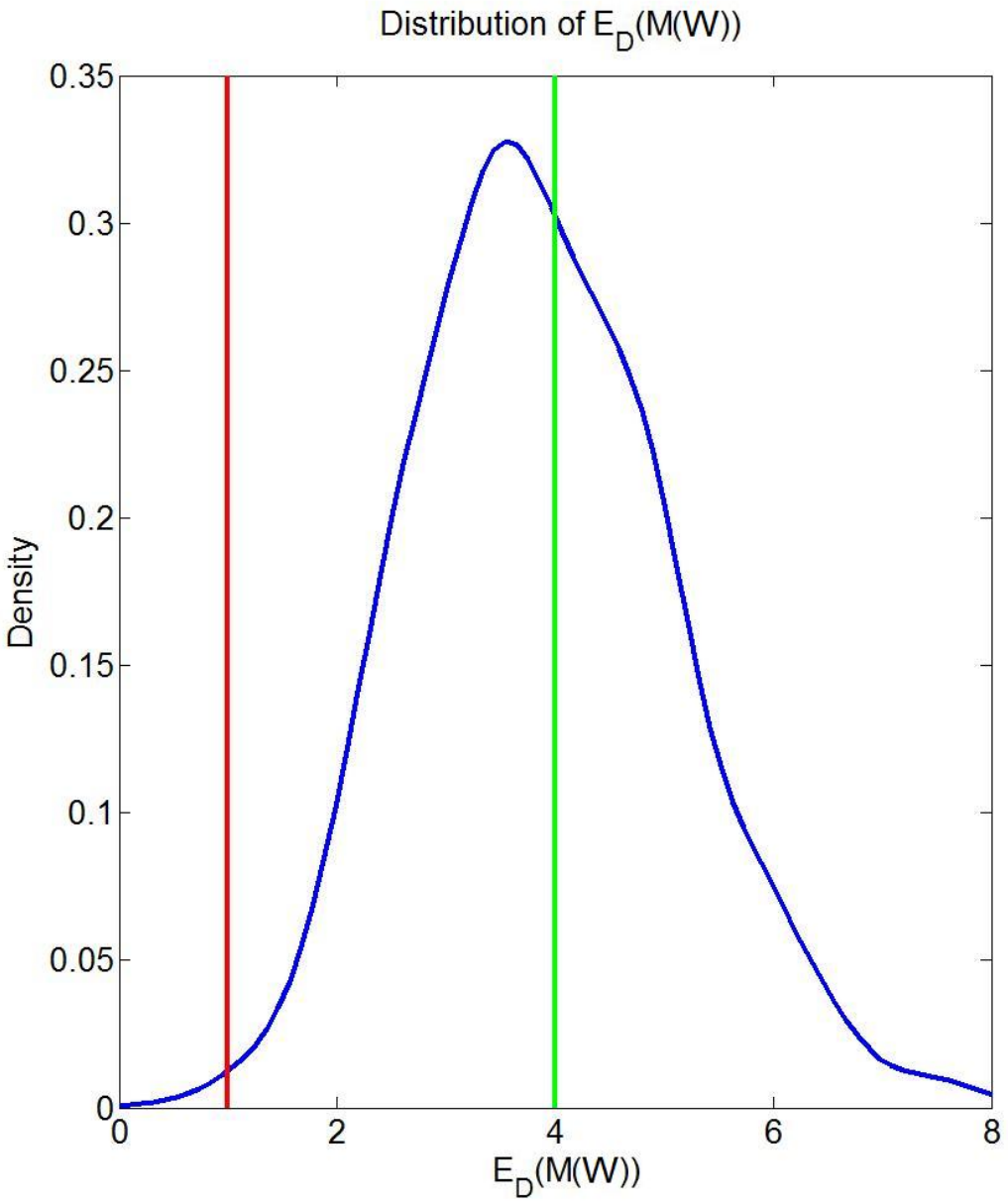

Figure 5.2: The range of $E_D(\mathcal{M}(W_X))$ using 50 sets of T=100. The red line shows the prior value for Σ_X. The green line shows the value of Σ_X for the "real" system.

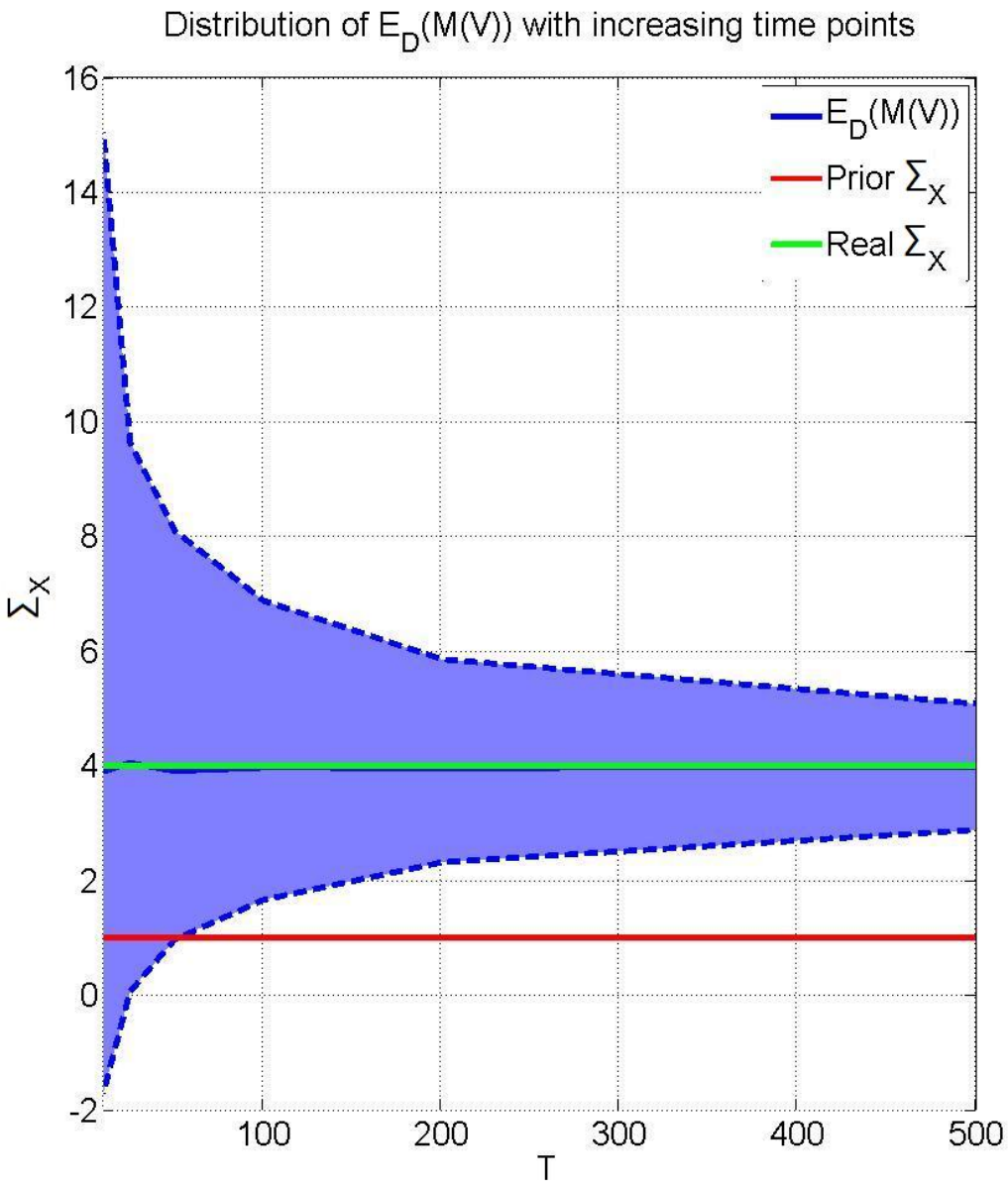

Figure 5.3: The uncertainty on $E_D(\mathcal{M}(W_X))$ as a function of number of observations. The shaded area shows the range of values predicted for $E_D(\mathcal{M}(W_X))$ over 50 synthetic data sets. Prediction of $E_D(\mathcal{M}(W_X))$, is centered at the correct value and prediction uncertainty decreases with increases data.

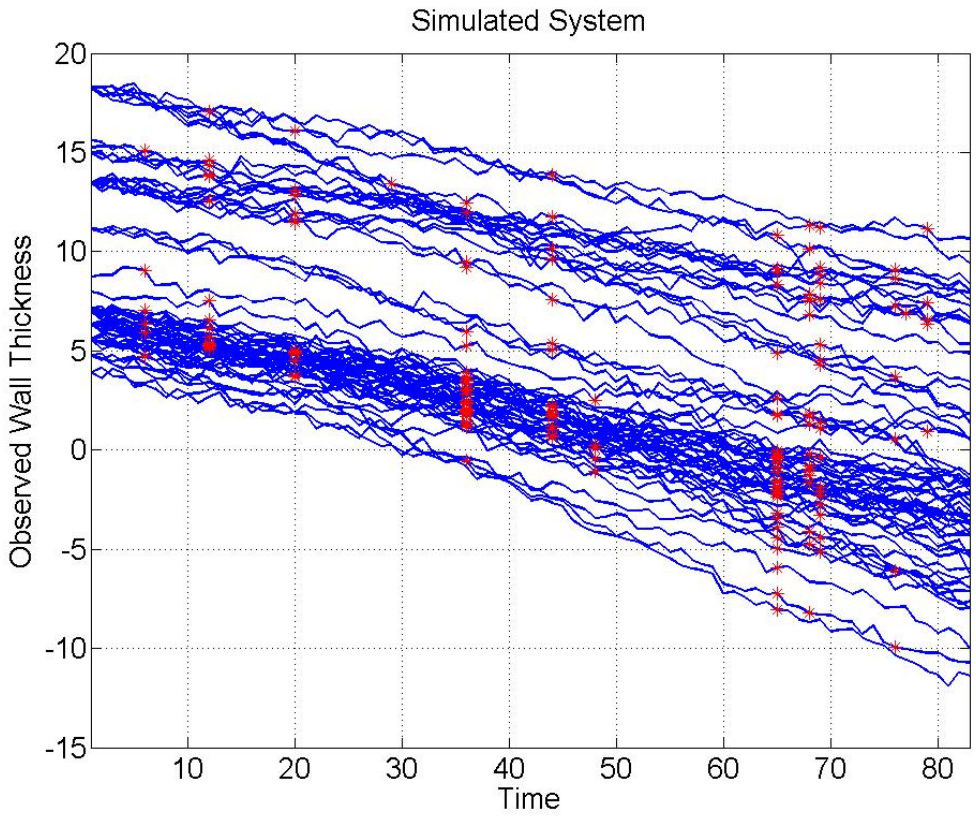

Figure 5.4: Synthetic data: observed wall thickness as a function of time. The observations of the systems are show with red points.

Updating Variances

To update the variances as described in section 5.5.3, using equation 5.5.28 we compute:

$$D_{ct} = (k_i Y_c^{(l_i)} - l_i Y_c^{(k_i)})^2$$

To compute D_{ct} we need at least 3 observations on a particular component, 42 of the 64 components have 3 or more inspections. This gives us 58 observed differences, D_{ct}. Our **prior** variance specification is given as:

$$\Sigma_Y = 0.04 \qquad \Sigma_X = 0.04 \qquad \Sigma_\alpha = 0.0001$$

The real system has variance:

$$\Sigma_Y = 0.04 \qquad \Sigma_X = 0.01 \qquad \Sigma_\alpha = 0.0001$$

with all variances the same except for the specification of the wall thickness variance. We specify 4th order quantities thus:

$$\Phi_{W_Y} = 0.001 \qquad \Phi_{W_X} = 0.001 \qquad \Phi_{W_\alpha} = 0.0001$$
$$\Xi_{W_Y} = 0.0005 \qquad \Xi_{W_X} = 0.0005 \qquad \Xi_{W_\alpha} = 0.00005.$$

These 4th order values chosen relative to the size of the prior variance.

Figure 5.5 shows the prior point-wise prior for the differences in this case. The expected value of $\mathrm{Dr}(Y_{ct})$ is 1, shown as a horizontal line. The horizontal line corresponding to $|1 - \mathrm{Dr}(Y_{ct})| = 3\sqrt{2}$ the $3\sigma_{\mathrm{Dr}Y}$ rule under normality, serving as a warning limit for unusually large values of discrepancy. The few points are over this warning limit are not huge suggesting the prior specification is reasonable.

Figure 5.6 shows the distribution of the adjustment, $E_D(\mathcal{M}(W_X))$ in this case. We see the adjusted expectation has moved in the right direction from the prior. One potential issue is that in cases where the adjustment is too big, $E_D(\mathcal{M}(W_X))$ can become negative, which for adjusting expectations about variances is an undesirable outcome. Getting negative variances can serve as a diagnostic warning of problems with the model.

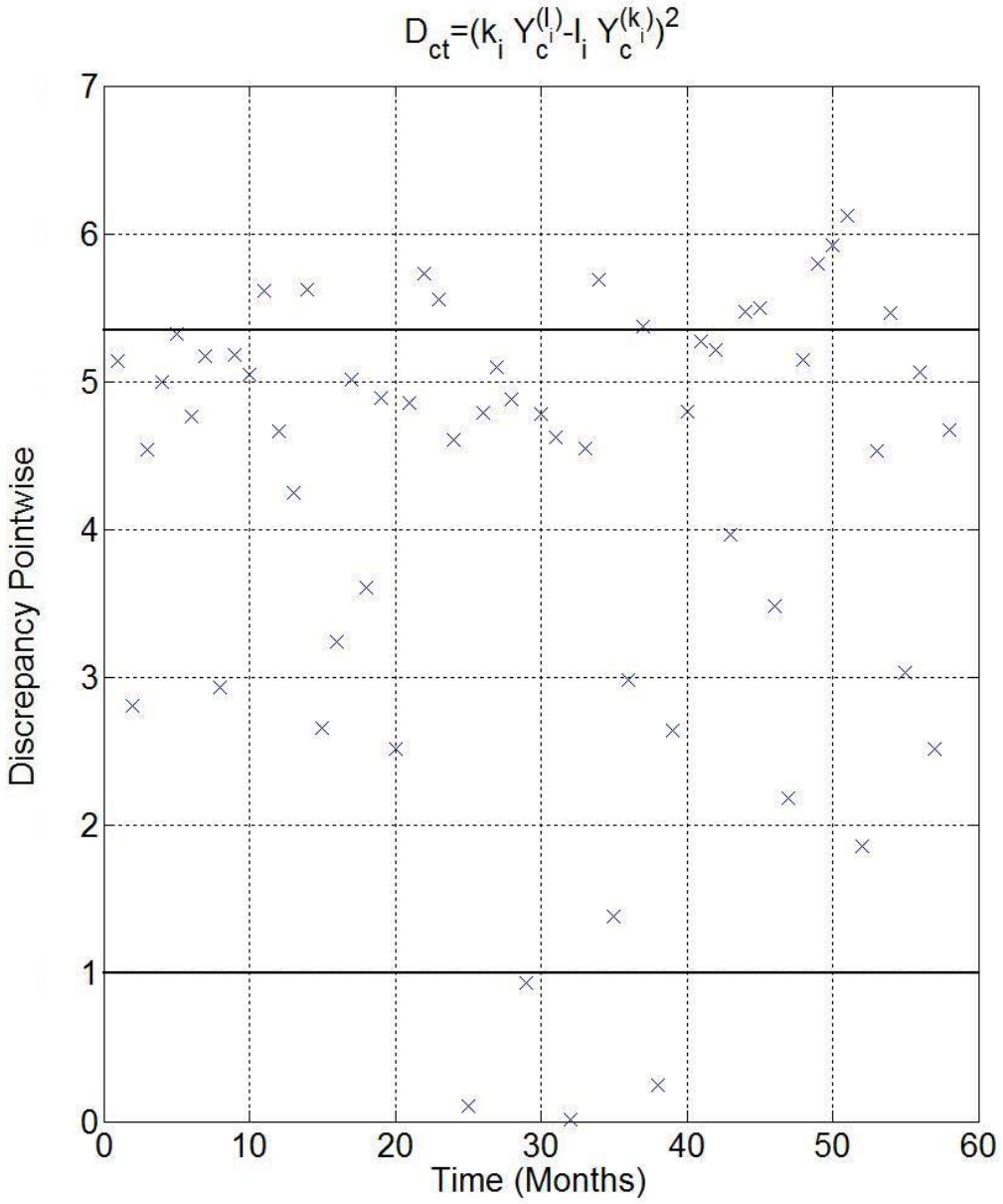

Figure 5.5: Point-wise Discrepancy for D_{ct}. The expected value of $\mathrm{Dis}(Y_{ct})$ is 1, shown as a horizontal line. The horizontal line corresponding to $|1-\mathrm{Dis}(Y_{ct})| = 3\sqrt{2}$ the $3\sigma_{\mathrm{Dr}Y}$ rule under normality, serving as a warning limit for unusually large values of discrepancy. A few points are over this warning limit are not huge suggesting the prior specification is reasonable

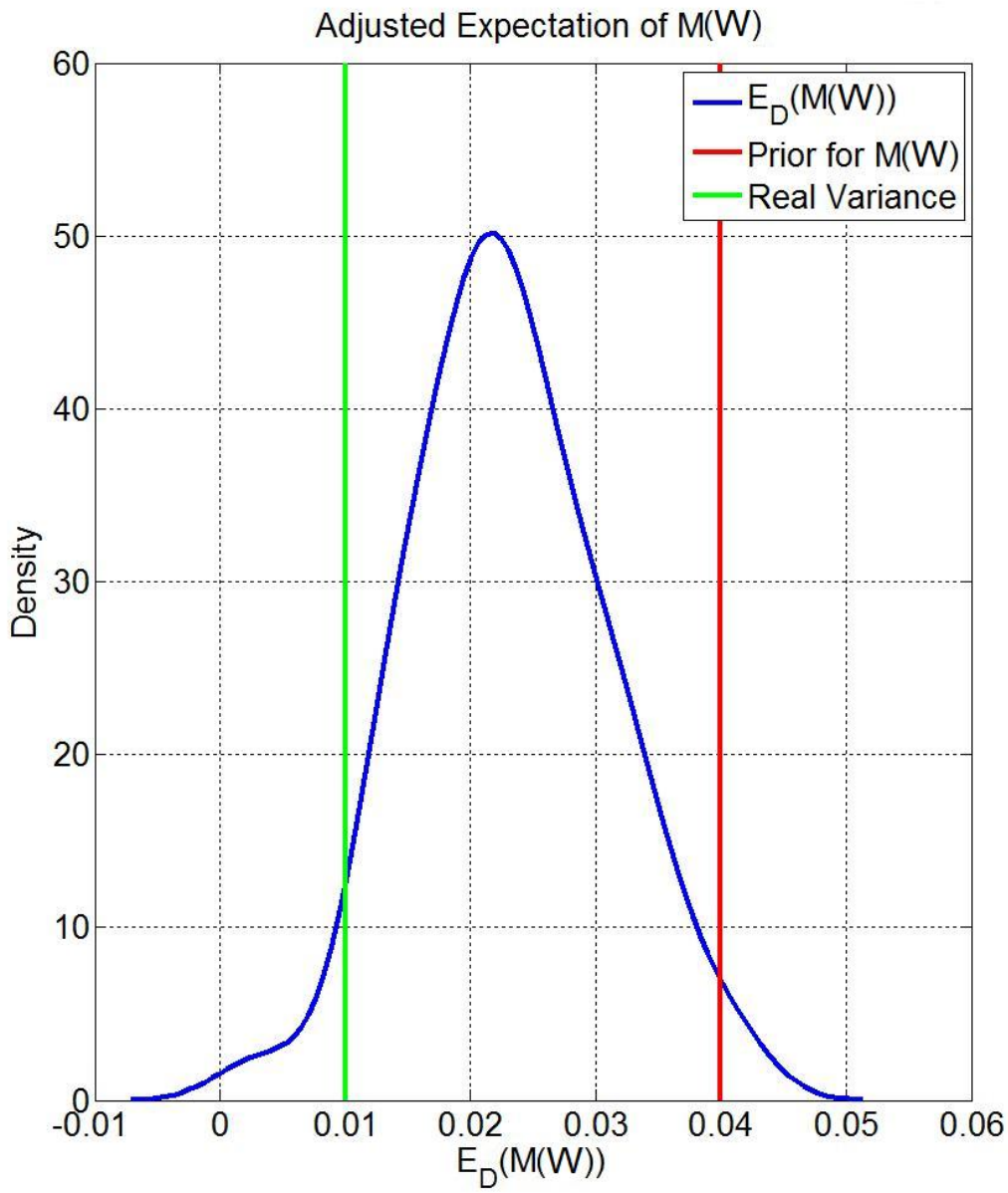

Figure 5.6: The range of $E_D(\mathcal{M}(W_X))$ using 50 sets of synthetic data. The red line shows the prior value for Σ_X. The green line shows the value of Σ_X for the "real" system.

Chapter 6

Bayes linear variance learning for the complete model with known local variance parameters

We now build on the ideas used in chapter 5 for a more general class of models; the model developed in chapter 3. We will use Bayes linear adjustment to update our beliefs about the global effects variances. As for variance learning in the case of the linear growth DLM, we need to still be able to handle the irregularly spaced, incomplete short time series data. The complete model has the added complication of the non-linear observation equation.

We will first consider the problems raised by a non linear observation equation, before extending ideas from the previous chapter to the more general setting. So we will consider updating the system variances for the global effects model in the cases of full inspections using linear combinations of observations and second order exchangeability assumptions. We will then extend this to irregular or partial inspections, before looking at an example using the oil platform application.

6.1 Non linear observation equation

The global effects model, section 3.1.2, is a general form DLM over a parameter space of N parameters with observations made on C components. We may carry

out variance learning for the global effects terms using a similar approach to that taken in section 5.3. However, we don't directly observe the global effects terms. Instead, we observe a potentially non linear function, f_l, which summarises the surface behaviour of the component. We can re-express the observation equation, i.e. from equation 3.3.11:

$$\begin{aligned} Y_{ct} &= f_l\left(F_c\Theta_t + r_{lt} + \epsilon_{Ylct}\right) \\ &= (F^*\Theta_t)_c + f_l\left(r_{lt} + \epsilon_{Ylct}\right) \\ &= F_c^*\Theta_t + M_{tc} \end{aligned}$$

where $M_{tc} = f_l\left(r_{lt} + \epsilon_{Ylct}\right)$ and F_c^* is the $1 \times N$ vector, for the cth row of the $C \times N$ dynamic regression matrix F^*.

We wish to find linear combinations of observations which eliminate model parameters, leaving expressions involving simply the residual structure using a similar approach to the linear growth model in section 5.3.3. We can then use second order exchangeability assumptions to update beliefs about the error structures and learn about underlying variance parameters.

However, we cannot find expressions which eliminate the effects of the local effects and measurement error terms inside the observation equation M_{tc}. This is because the observation equation is non linear and we have no expression for its evolution in time. Instead, assuming known local variance parameters we can simulate relevant information about it, e.g. $E(M_{tc})$, $\mathrm{Var}(M_{tc})$ and $\mathrm{Cov}(M_{tc}, M_{tc'})$, as explained in section 4.6. A method is presented for Bayes linear variance learning for the general variance parameters with known local variance parameter. In chapter 7 this restriction is then lifted to allow learning about all variance parameters within the model.

6.2 Bayes linear variance learning for corrosion model

Consider learning about the variances within the corrosion model section, 3.5. This model has a non linear, minimum observation equation 3.5.20. We seek expressions

for squared residuals ϵ^2_{Xtc} from sample data corresponding to partial system inspections. The general effects DLM part of the model is a linear growth DLM. We can take linear combinations of observations to isolate expressions for the residuals even when observations of the system are irregularly spaced in time. Using expressions for ϵ_{Xtc} we adjust our beliefs about the population mean variance using observed data and assumed known values for the the local corrosion variances as explained below.

6.2.1 Linear combinations of observations

Firstly we shall consider the case of full inspections. We consider differences of observations in time; using equations 3.5.16 and 3.5.20.

Let $Y_c^{(1)}$ and $M_c^{(1)}$ be the 1 step differences:

$$Y_c^{(1)} = Y_{ct} - Y_{c(t-1)}$$
$$M_c^{(1)} = M_{ct} - M_{c(t-1)}$$

In the case of the corrosion model using equation 3.5.20:

$$M_{ct} = \min_l \left(r_{lct} + \epsilon_{Ylct}\right).$$

The minima term, M_{ct} is independent of the general effects model. Consequently assuming known local effects variance parameters and measurement error the local effects model can be simulated as discussed in section 4.6. Any quantities of interest about M_{ct}, ($E(M_{ct})$, $\mathrm{Var}(M_{ct})$, etc.) can be then be estimated empirically from the simulation.

Consider the one time step difference:

$$\begin{aligned} Y_c^{(1)} &= Y_{ct} - Y_{c(t-1)} \\ &= X_{ct} - X_{c(t-1)} + M_{ct} - M_{c(t-1)} \\ &= X_{c(t-1)} + \alpha_{ct} + \epsilon_{Xct} - X_{c(t-1)} + M_c^{(1)} \\ &= \alpha_{ct} + \epsilon_{Xct} + M_c^{(1)} \\ &= \alpha_{c(t-1)} + \epsilon_{\alpha ct} + \epsilon_{Xct} + M_c^{(1)} \end{aligned}$$

and the two step differences:

$$\begin{aligned} Y_c^{(2)} &= Y_{ct} - Y_{c(t-2)} \\ &= X_{ct} - X_{c(t-2)} + M_c^{(2)} \\ &= X_{c(t-1)} + \alpha_{ct} + \epsilon_{Xct} - X_{c(t-2)} + M_c^{(2)} \\ &= X_{c(t-2)} + 2\alpha_{c(t-1)} + \epsilon_{\alpha c(t-1)} + \epsilon_{\alpha ct} + \epsilon_{Xc(t-1)} + \epsilon_{Xct} - X_{c(t-2)} + M_c^{(2)} \\ &= 2\alpha_{c(t-1)} + \epsilon_{\alpha ct} + \epsilon_{Xc(t-1)} + \epsilon_{Xct} + M_c^{(2)} \end{aligned}$$

then the linear combination below isolates the residual error structures from the the general corrosion model:

$$Y_c^{(2)} - 2Y_c^{(1)} = -\epsilon_{\alpha tc} + \epsilon_{Xt-1c} - \epsilon_{Xtc} + M_c^{(2)} - 2M_c^{(1)}$$

This leaves terms involving the residual structure of the DLM and terms involving the minimum. For known fixed local variance, the effect of terms involving the minima can be simulated.

6.2.2 Adjusting beliefs using exchangeability

Let:

$$D_{ct} = (Y_c^{(2)} - 2Y_c^{(1)})^2$$

and:

$$D = \begin{pmatrix} D_{11} \\ \vdots \\ D_{CT} \end{pmatrix}.$$

Using second order exchangeability assumptions described in section 3.5.1 and taking the square of this linear combination and taking expectations we get:

$$\begin{aligned} E(D_{ct}) &= E\left[(Y_c^{(2)} - 2Y_c^{(1)})^2\right] \\ &= E\left[(\epsilon_{\alpha tc} + \epsilon_{Xt-1c} - \epsilon_{Xtc} + M_c^{(2)} - 2M_c^{(1)})^2\right] \\ &= E(\epsilon_{\alpha tc}^2) + E(\epsilon_{Xt-1c}^2) + E(\epsilon_{Xt-1c}^2) + E((M_c^{(2)} - 2M_c^{(1)})^2) \\ &= E(\mathcal{M}(W_\alpha) + 2\mathcal{M}(W_X)) + E((M_c^{(2)} - 2M_c^{(1)})^2) \\ &= \Sigma_{W_\alpha} + 2\Sigma_{W_X} + E((M_c^{(2)} - 2M_c^{(1)})^2) \end{aligned}$$

For known local corrosion variance parameters we can simulate the local corrosion terms involving the minimum function $E((M_c^{(2)} - 2M_c^{(1)})^2)$.

Therefore to learn about general corrosion variances we can compute the adjusted expectation, $E_D(\mathcal{M}(W_X))$:

$$\begin{aligned} E_D(\mathcal{M}(W_X)) &= E(\mathcal{M}(W_X)) + \text{Cov}(\mathcal{M}(W_X), D)(\text{Var}(D))^{-1}(D - E(D)) \\ &= \Sigma_{W_X} + 1_n^T(\Xi_{W_X})(\text{Var}(D))^{-1}(D - E(D)) \end{aligned}$$

where we can use simulations to find estimates for $\text{Var}(D))^{-1}$. Calculations for $\text{Cov}(\mathcal{M}(W_X), D)$ is discussed for irregularly spaced observations in the appendix B.1.

6.2.3 Irregular time steps

In section 5.5 we considered the problem of learning about the linear growth DLM model in the case where we don't have full inspections. The general corrosion model from equation 3.5.16 is given by:

$$\begin{aligned} X_{ct} &= X_{c(t-1)} + \alpha_{ct} + \epsilon_{Xct} \\ \alpha_{ct} &= \alpha_{c(t-1)} + \epsilon_{\alpha ct} \end{aligned} \tag{6.2.1}$$

The general corrosion DLM can be rewritten to tell us about time steps longer than one step:

$$\begin{pmatrix} X \\ \alpha \end{pmatrix}_{tc} = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ \alpha \end{pmatrix}_{t-kc} + \sum_{i=0}^{k-1} \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{t-ic}$$

6.2.4 Linear combinations of observations

We want to isolate the squared residual error terms and use them to update our beliefs about the underlying variance parameters. As in section 5.5.3, given observations for components c:

$$\{Y_{ct_1}, Y_{ct_2}, \ldots, Y_{ct_{T_c}}\}$$

at times:

$$\{t_1, t_2, \ldots, t_{T_c}\}$$

then let k_i be time between observations, t_i and t_{i-1}; let l_i be time between observations, t_i, and t_{i-2}:

$$k_i = t_i - t_{i-1} \qquad\qquad l_i = t_i - t_{i-2}$$

where T_c is the total number of observations of component c. As in equation 5.5.28 we take:

$$D_{ct} = l_i Y_c^{(k_i)} - k_i Y_c^{(l_i)}$$

where:

$$E(D_{ct}) = \frac{k_i l_i (k_i - l_i)(2k_i^2 - 2k_i l_i - 1)}{6}\Sigma_\alpha + k_i l_i (l_i - k_i)\Sigma_X + E\left(l_i (M_c^{(k_i)} - k_i M_c^{(l_i)})^2\right)$$

The terms involving $E[M_c^{(k_i)2}], E[M_c^{(l_i)2}], E[M_c^{(l_i)} M_c^{(k_i)}]$ can be estimated from simulations for assumed known values of the local corrosion error variance.

To update our beliefs about the variance we compute $E_D(\mathcal{M}(W_X))$:

$$\begin{aligned} E_D(\mathcal{M}(W_X)) &= E(\mathcal{M}(W_X)) + \mathrm{Cov}(\mathcal{M}(W_X), D)(\mathrm{Var}(D))^{-1}(D - E(D)) \\ &= \Sigma_{W_X} + 1_C^T(\Xi_{W_X})(\mathrm{Var}(D))^{-1}(D - E(D)) \end{aligned}$$

The calculation of $\mathrm{Cov}(\mathcal{M}(W_X), D)$ is given in the appendix, B.1 and $\mathrm{Var}(D)$ is estimated using simulation.

6.3 Bayes linear variance learning for the complete model

6.3.1 Linear combinations of observations

In the case of full inspections, we take differences of observations in time, then use equations 3.1.1, 3.2.7, 3.3.10 and 3.3.11.

Let $Y_c^{(1)}$ and $M_c^{(1)}$ be the ith step differences:

$$Y_{ct}^{(1)} = Y_{ct} - Y_{c(t-1)}$$
$$M_{ct}^{(1)} = M_{ct} - M_{c(t-1)}$$

We first consider the one time step difference:

$$\begin{aligned} Y_{ct}^{(1)} &= Y_{ct} - Y_{c(t-1)} \\ &= F_c^*\Theta_t - F_c^*\Theta_{t-1} + M_{ct} - M_{c(t-1)} \\ &= F_c^*(\Theta_t - \Theta_{t-1}) + M_{ct}^{(1)} \\ &= F_c^*\big(G\Theta_{t-1} - \Theta_{t-1} + \epsilon_{\Theta t}\big) + M_{ct}^{(1)} \\ &= F_c^*\big((G - I_N)\Theta_{t-1} + \epsilon_{\Theta t}\big) + M_{ct}^{(1)} \\ &= F_c^*\big((G - I_N)G\Theta_{t-2} + \epsilon_{\Theta t} + (G - I_N)\epsilon_{\Theta t-1}\big) + M_{ct}^{(1)} \end{aligned} \tag{6.3.2}$$

and the two step differences:

$$\begin{aligned} Y_{ct}^{(2)} &= Y_{ct} - Y_{c(t-2)} \\ &= F_c^*\Theta_t - F_c^*\Theta_{t-2} + M_{ct}^{(2)} \\ &= F_c^*\big(G\Theta_{t-1} - \Theta_{t-2} + \epsilon_{\Theta t}\big) + M_{ct}^{(2)} \\ &= F_c^*\big(G^2\Theta_{t-2} - \Theta_{t-2} + \epsilon_{\Theta t} + G\epsilon_{\Theta t-1}\big) + M_{ct}^{(2)} \\ &= F_c^*\big((G^2 - I_N)\Theta_{t-2} + \epsilon_{\Theta t} + G\epsilon_{\Theta t-1}\big) + M_{ct}^{(2)} \end{aligned} \tag{6.3.3}$$

We now want a linear combination which removes the system state effects, Θ_t leaving an expression involving the residual structure and terms involving M_{ct} which can be simulated.

Lemma 6.3.1 Consider the linear combination:

$$G^T(G - I_N)^T(F_c^*)^T Y_{ct}^{(2)} - (G^2 - I_N)^T(F_c^*)^T Y_{ct}^{(1)} \tag{6.3.4}$$

If the $N \times N$ matrix:

$$(G^2 - I_N)^T(F_c^*)^T F_c^*(G - I_N)G \tag{6.3.5}$$

is **symmetric** then the linear combination in equation 6.3.4, removes the system state effects, Θ_t and leaves terms involving only the residual structure as required for variance learning.

Proof:

Using equations 6.3.2 and 6.3.3 the linear combination in equation 6.3.4 gives:

$$\begin{aligned} G^T(G-I_N)^T(F_c^*)^T Y_{ct}^{(2)} &= G^T(G-I_N)^T(F_c^*)^T F_c^*(G^2-I_N)\Theta_{t-2} \\ &\quad + G^T(G-I_N)^T(F_c^*)^T M_{ct}^{(2)} \qquad (6.3.6) \\ (G^2-I_N)^T(F_c^*)^T Y_{ct}^{(1)} &= (G^2-I_N)^T(F_c^*)^T F_c^*(G-I_N)G\Theta_{t-2} \\ &\quad + (G^2-I_N)^T(F_c^*)^T M_{ct}^{(1)} \qquad (6.3.7) \end{aligned}$$

note that the coefficient of Θ_{t-2} in 6.3.6 is the transpose of the coefficient of Θ_{t-2} in 6.3.7. Therefore if the matrix in equation 6.3.5. □

Lemma 6.3.1 holds for the linear growth DLM, in this case:

$$G = \begin{pmatrix} I_C & I_C \\ 0_C & I_C \end{pmatrix} \qquad\qquad F' = \begin{pmatrix} I_C & 0_C \end{pmatrix}$$

and so:

$$\begin{aligned} &(G^2-I_N)^T(F')^T F'(G-I_N)G = \\ &= \left(\left(\begin{pmatrix} I_C & I_C \\ 0_C & I_C \end{pmatrix}^2 - I_N\right)^T \begin{pmatrix} I_C \\ 0_C \end{pmatrix} \begin{pmatrix} I_C & 0_C \end{pmatrix}\right. \\ &\quad \left.\times \left(\begin{pmatrix} I_C & I_C \\ 0_C & I_C \end{pmatrix} - I_N\right) \begin{pmatrix} I_C & I_C \\ 0_C & I_C \end{pmatrix}\right) \\ &= \begin{pmatrix} 0_C & 2I_C \\ 0_C & 0_C \end{pmatrix}^T \begin{pmatrix} I_C & 0_C \\ 0_C & 0_C \end{pmatrix} \begin{pmatrix} 0_C & I_C \\ 0_C & 0_C \end{pmatrix} \\ &= \begin{pmatrix} 0_C & 0_C \\ 0_C & 2I_C \end{pmatrix} \end{aligned}$$

this is symmetric and so lemma 6.3.1 holds, and the linear combination in equation, 6.3.4 allows us to get an expression involving only the residual structure.

Lemma 6.3.1 also hold in the case where:

$$G = \begin{pmatrix} I_C & I_C \\ I_C & I_C \end{pmatrix}$$

This linear combination of observations allows us to carry out variance learning using the same approach as in section 6.2. For model forms which this lemma does

not apply it is possible to find other similar linear combinations to get an expression involving only the residual structure or which assess different parts of the residual structure.

6.3.2 Adjusting beliefs using exchangeability

From section 3.1.3 we have exchangeability assumptions which express our beliefs about the relationship of the residuals within the global effects model. We consider the residuals exchangeable in time from equation 3.1.2:

$$\epsilon_{\Theta nt} = \mathcal{M}(\epsilon_{\Theta n}) + \mathcal{R}_t(\epsilon_{\Theta n}),$$

where from equation 3.2.9.

$$\mathcal{M}(\epsilon_{\Theta n}) = 0$$

we also assume the squared residuals are exchangeable in time from equation 3.1.5:

$$[\epsilon_{\Theta nt}]^2 = [\mathcal{R}_t(\epsilon_{\Theta n})]^2 = V_{nt} = \mathcal{M}(V_n) + \mathcal{R}_t(V_n)$$

in the same way as in section 5.4, we also make assumptions about the residual structure across components for parts of the parameter space, $\mathcal{M}(\epsilon_{\Theta n})$, i.e. from equation 3.1.6:

$$\mathcal{M}(V_p) = W_p = \mathcal{M}(W) + \mathcal{R}_p(W)$$

and so we have that:

$$\epsilon^2_{\Theta nt} = \mathcal{M}(W) + \mathcal{R}_p(W) + \mathcal{R}_t(V_p)$$

and using section 2.4 it is sufficient to update our beliefs about $\mathcal{M}(W)$ to learn about the variance.

Given observational data $\{Y_{ct}\}$ we can construct data vector D of squared linear combinations where:

$$D = \begin{pmatrix} D_{11} \\ D_{12} \\ \vdots \\ D_{1T-2} \\ \vdots \\ D_{CT-2} \end{pmatrix}$$

and:

$$D_{ct} = \left[(G - I_N)G(F_c^*)^T Y_{ct}^{(2)} - (G^2 - I_N)(F_c^*)^T Y_{ct}^{(1)}\right]^2$$

Under the assumption, that for a given model form, G and F', the expression in lemma 6.3.1 is symmetric.

To learn about general effects variances we can compute the adjusted expectation $E_D(\mathcal{M}(W))$:

$$E_D(\mathcal{M}(W)) = E(\mathcal{M}(W)) + \text{Cov}(\mathcal{M}(W), D)(\text{Var}(D))^{-1}(D - E(D))$$

Given specific model forms we could either find expressions for $E(D)$, $\text{Var}(D)$ and $\text{Cov}(\mathcal{M}(W), D)$ or in cases where they are difficult to compute in closed form use simulation. This approach can also be extended to irregularly spaced observations in the same way as we have shown for the corrosion model.

6.4 Example: Updating the variances for offshore platform application

We continue the full-scale offshore platform example (previously discussed in sections 1.4.1, 3.5 and 4.9, 5.7.1 and 5.7.2).

In the example in the previous chapter, section 5.7.2 we considered learning about wall thickness variances for a linear growth DLM in the case of synthetic data (simulated using the prior mean specification as used in for the offshore platform example 4.9). In this example we will update beliefs about variances using the synthetic data set for component minimum wall thickness, simulated using the corrosion model described in section 3.5.

6.4.1 Prior specification

The mean prior specification is the same as used in the offshore platform example in section 4.9. We set the number of locations as 20. Our **prior** variance specification is given as:

$$\Sigma_Y = 0.167 \qquad \Sigma_X = 0.05 \qquad \Sigma_r = 0.1 \qquad \Sigma_\alpha = 0.001$$

The "real" system has variance:

$$\Sigma_Y = 0.167 \qquad \Sigma_X = 0.1 \qquad \Sigma_r = 0.1 \qquad \Sigma_\alpha = 0.001$$

The prior wall thickness variance has been underestimated. As in the example in the previous chapter, section 5.7.1 we will investigate how well we can recover the "real" system variances. Figure 6.1 shows the point wise prior discrepancy of one realisation of the synthetic data. A large number of simulations are done as described in section 4.8.

6.4.2 Updating Variances

To illustrate Bayes Linear Variance learning, consider estimating Σ_{W_X} using equations above in this chapter, in a simulated case for which the actual value is set at 0.1. For each of 50 independent realisations of the inspection data sample, we generate D which is used to evaluate $E_D(\mathcal{M}(W_X))$ which is an estimate for Σ_{W_X}.

The distribution of the adjusted variance, $E_D(\mathcal{M}(W_X))$ is given in figure 6.2.e see the distribution of the $E_D(\mathcal{M}(W_X))$ covers the "real" value. We find that the mean value of $E_D(\mathcal{M}(W_X))$ is 0.0994, and that the 5% and 95% values are respectively 0.0872 and 0.1128, consistent with the real value. However the mode of the distribution is slightly to the left of the of the real value, probably the effect of the low prior.

Given known local variance terms we can update our beliefs about variances in the general part of the model. In practice the local corrosion error variance is unknown. The selection of optimal combination of local and general variance estimates consistent with the data is discussed in chapter 7.

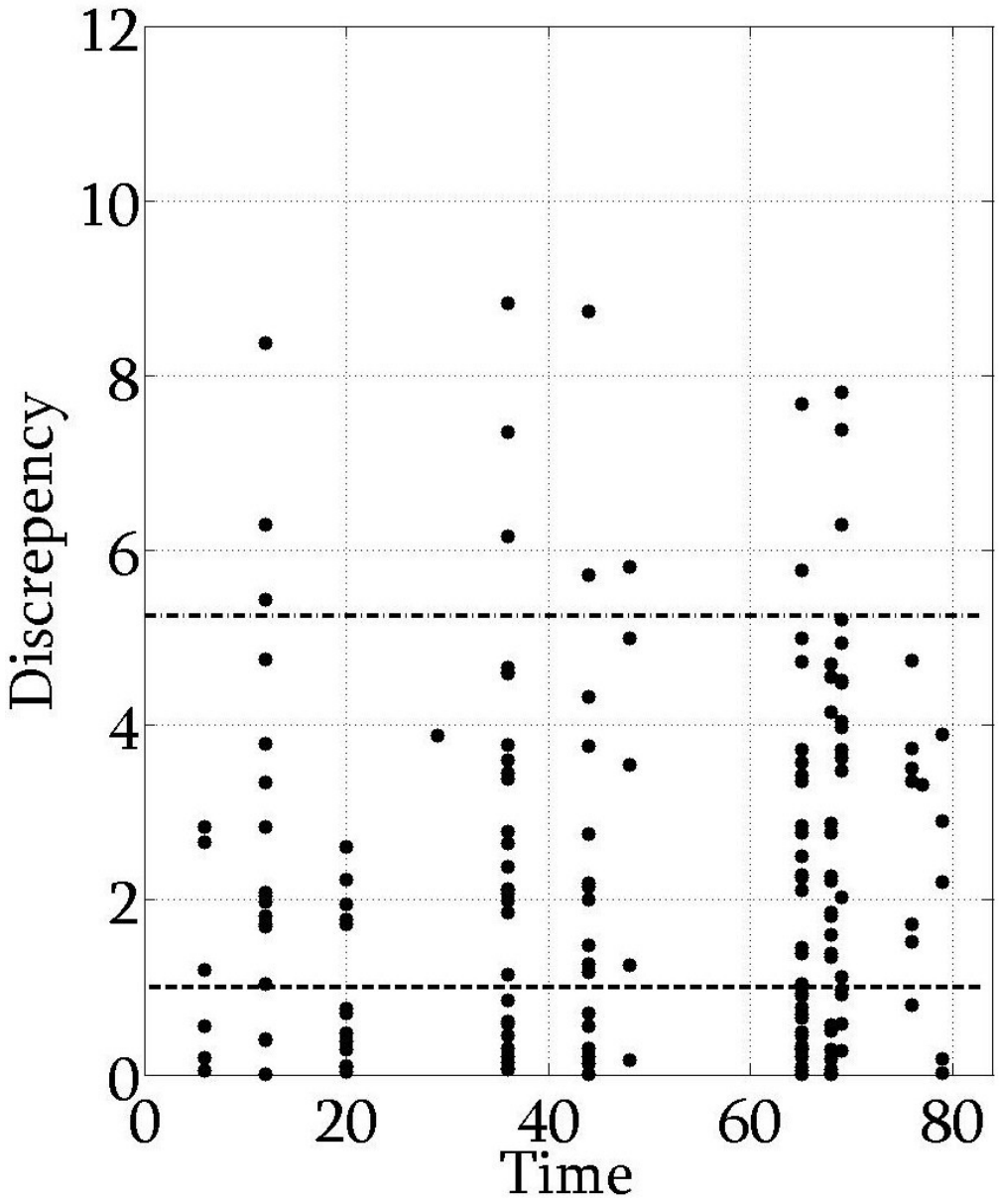

Figure 6.1: The point wise discrepancy of one realisation of the synthetic inspection data. The expected value of $\mathrm{Dis}(Y_{ct})$ is 1, shown as a horizontal line. Also shown is the horizontal line corresponding to $|1 - \mathrm{Dis}(Y_{ct})| = 3\sqrt{2}$ the $3\sigma_{\mathrm{Dr}Y}$ rule under normality, serving as a warning limit for unusually large values of discrepancy.

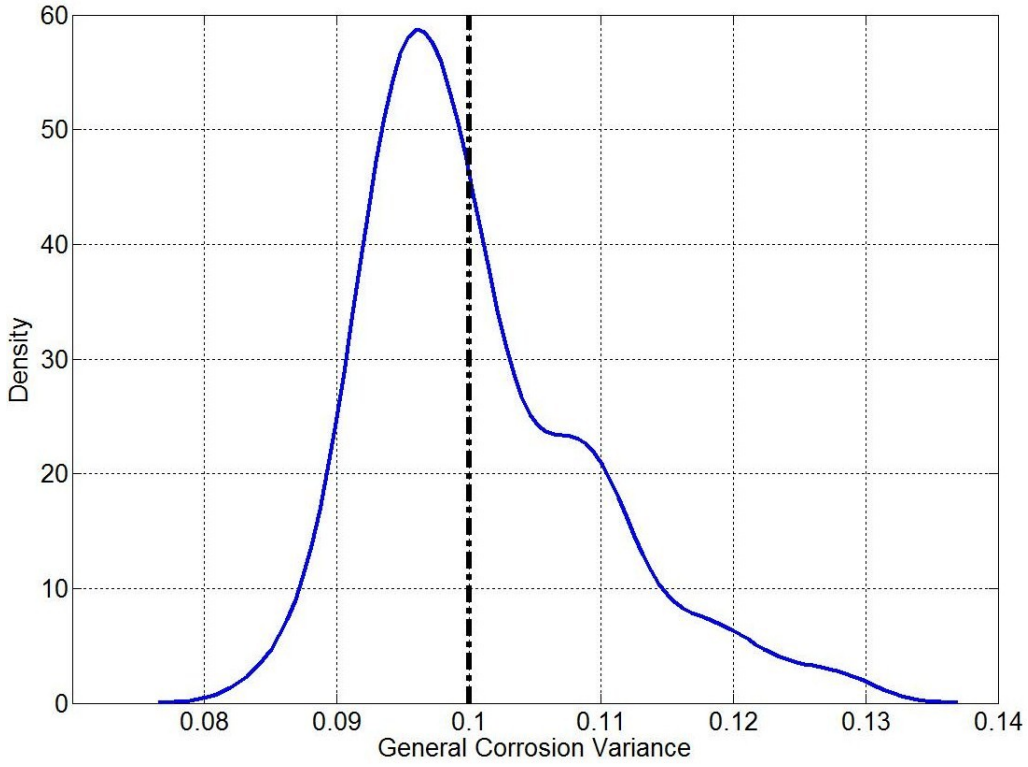

Figure 6.2: Kernel density estimate of adjusted variance $E_D(\mathcal{M}(W_X))$ using 50 sets of synthetic data for offshore platform example. The "real" general corrosion variance is $\Sigma_{W_X} = 0.1$ shown by the vertical line. We see the distribution of the $E_D(\mathcal{M}(W_X))$ covers the "real" value. However the mode of the distribution is slightly to the left of the of the real value, probably the effect of the low prior.

Chapter 7

Mahalanobis variance learning

In the preceding chapters 5 and 6 it has been shown that we can use exchangeability assumptions and Bayes linear adjustments to estimate variances in the global effects model. However we have not considered learning about the local effects variances. Unlike the global effects model variance, we cannot find linear combinations of observations which eliminate parameter system state effects. In this chapter we will describe a different approach using the Mahalanobis distance to estimate local variance parameters. The approach is similar to the "history matching" discussed in Craig et al. [1997], who try to find inputs to a computer simulator that closely match historical data in the context of oil reservoirs. Note that this method is particularly appropriate when we are unsure whether there are *any* choices of parameters which will give a good fit to the data, as the approach combines parameter estimation with goodness of fir assessment.

7.1 Bayes linear adjustment for local effects

In section 5.6 we described the use of Bayes linear updating to adjust our beliefs for the global effects evolution error variances in the case of known local effects error variances. Updating the global effects uses the linear structure of DLM to update beliefs and get expressions for the residual structure. Ideally we would use a Bayes linear scheme to update the local corrosion variances directly. However the local

effects model is embedded within a non-linear observation equation 3.3.12:

$$Y_{ct} = (F^{*}\Theta_t)_c + f_l\left(r_{lct} + \epsilon_{Ylct}\right)$$

therefore we cannot use Bayes linear assessment directly. Suppose that we consider a new random quantity:

$$M_{lct} = f_l\left(r_{lct} + \epsilon_{Ylct}\right).$$

Given prior specification for M_{lct} and suitable exchangeability assumptions we could use Bayes linear variance learning to update beliefs about $\mathrm{Var}(M)$. However this would tell us nothing directly about variance parameters within the observation equation. Linear combinations of observations cannot be used to provide direct estimates for local corrosion evolution variance since the observation equation is non-linear.

We instead adopt a straightforward fitting procedure to estimate optimal combinations of local and general corrosion error variances which are most consistent with observational data. The approach adopts a Mahalanobis distance fitting criterion, exploiting the covariance structure to aid variance learning.

7.2 Mahalanobis learning: normal distribution

We now consider how to carry out parameter estimation using the Mahalanobis distance. The approach is similar to a profile likelihood, however we can use it in cases where we don't have a likelihood . Properties of the Mahalanobis distance have been discussed in previous section 4.7.2. We begin with a simple example, intended to introduce the ideas involved in Mahalanobis learning. This simple case should aid in understanding of the less tractable cases.

Using the Mahalanobis distance we will estimate the parameters in a normal distribution.

We observe a set of, n, observations , $D_{[n]}$:

$$D_{[n]} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

where each $X_k \sim N(\mu, \sigma^2)$. Use the observations $D_{[n]}$ we estimate parameters μ and σ from two candidate sets of potential parameter values:

$$\mu = \{\mu_1, \ldots, \mu_i, \ldots, \mu_m\}$$
$$\sigma = \{\sigma_1, \ldots, \sigma_j, \ldots, \sigma_p\}$$

Given a particular choice of parameters, μ_i, σ_j we can then test its fitting performance to data by computing the Mahalanobis distance. The observational data, $D_{[n]}$, can be compared with each possible combination of μ and σ to determine the choice which best fit the data.

For a particular choice of μ_i, σ_j, from equation 4.7.7, the Mahalanobis distance, $\text{Dis}(D_{[n]})$, is given by:

$$\begin{aligned}\text{Dis}(D_{[n]}) &= (\mu_i - D_{[n]})^T \left[\text{diag}(\sigma_j^2)\right]^{-1} (\mu_i - D_{[n]}) \\ &= \sum_{k=1}^{n} \frac{(\mu_i - X_k)^2}{\sigma_j^2}\end{aligned}$$

and standardised Mahalanobis distance, $\text{Dr}(D_{[n]})$, is given as:

$$\text{Dr}(D_{[n]}) = \frac{1}{n} \sum_{k=1}^{n} \frac{(\mu_i - X_k)^2}{\sigma_j^2}$$

In following sections we will consider a running example using a vector,$D_{[100]}$, of $n = 100$ observations, of X, from a normal distribution;

$$X \sim N(10, 4).$$

7.2.1 Learning about μ

Firstly, treating σ as known we learn about μ given observations, $D_{[100]}$ and candidate set:

$$\mu = \{1, 1.5, 2, ..., 20\}.$$

Figure 7.1a shows standardised Mahalanobis distance,$\text{Dr}(D_{[100]})$, over the range of the candidate set, μ. The value of Dr at the minimum is ≈ 1, when μ near true value of 10, consistent with expected value of Dr.

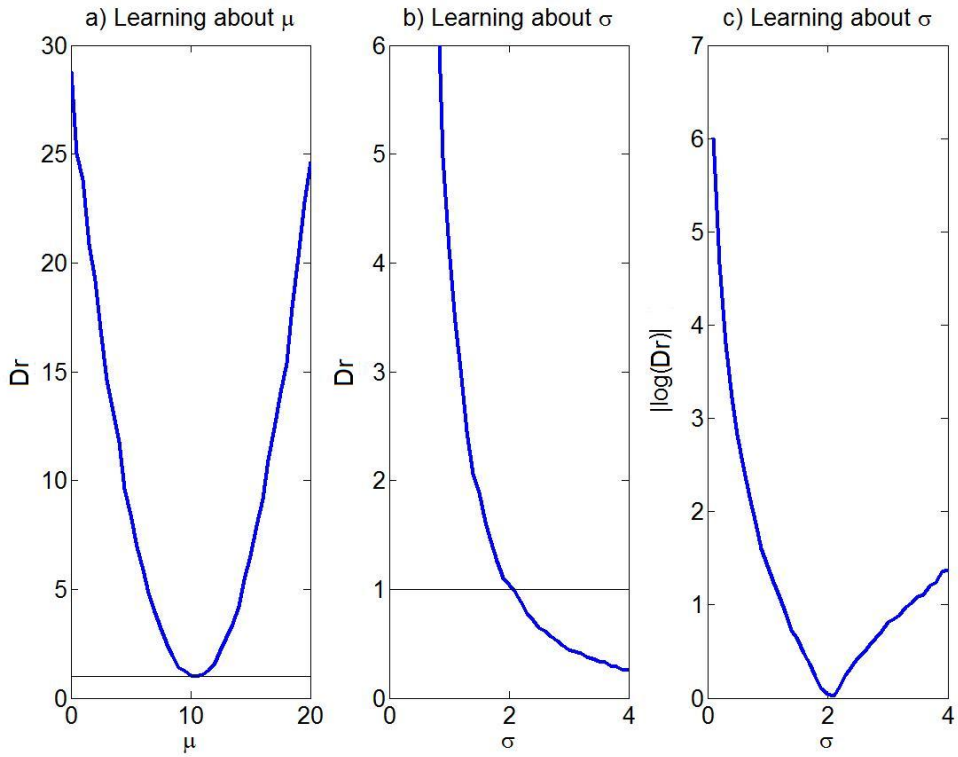

Figure 7.1: Mahalanobis learning for parameter in a normal distribution. a) shows $\mathrm{Dr}(D_{[100]})$ over the range of the candidate set for μ with horizontal line for the expected value of 1. b) shows the value of $\mathrm{Dr}(D_{[100]})$ over the candidate set for σ with horizontal line for the expected value of 1. c) shows |log(Dr)| over the range of σ. In the case of |log(Dr)| the expected value is 0 and the minimum is around the correct value.

7.2.2 Learning about σ

So now lets try and learn about the value of σ using data $D_{[100]}$. Using the same example but treating μ as known and fixed and we test the candidate set:

$$\sigma = \{0.1, 0.2, ..., 4\}.$$

Figure 7.1b shows the value of $\mathrm{Dr}(D_{[100]})$ over the candidate set σ:

$$\mathrm{Dis}(D_{[100]}) = \frac{1}{\sigma_j^2} \sum_{k=1}^{100} (\mu_i - X_k)^2$$

When searching over the variance parameter σ with fixed mean, μ_i this becomes:

$$\mathrm{Dis}(D_{[100]}) = \frac{\kappa}{\sigma_j^2} \qquad\qquad \kappa = \sum_{k=1}^{100} (\mu_i - X_k)^2$$

Consequently, for μ fixed:

$$\lim_{\sigma \to \infty} (\mathrm{Dis}(D_{[100]}) = 0.$$

This can be seen in figure 7.1b. At the real value $\sigma = 2$, $\mathrm{Dr} \approx 1$ the expected value. Regions where the Mahalanobis distance is small can either mean we have a good parameter estimate or very large uncertainty about that point. These regions are "not implausible" values but may still represent bad parameter choices for example very large variances.

Alternatively, figure 7.1c shows $|\log(\mathrm{Dr})|$ over the range of σ. Here we get a distinct minimum at $\sigma \approx 2$ and avoid the issues arising as $\mathrm{Dr} \to 0$.

7.2.3 Learning about μ and σ jointly

Consider learning about μ and σ jointly. Here we consider the grid of points formed by every combination of the candidate sets μ and σ. Figure 7.2a shows a heat-map where the range of "not implausible" values is everything not in the deep red. Regions where the Mahalanobis distance is large represent bad parameter choices and so can be ruled out as being "implausible".

This forms a wedge shape which will continue on to $\sigma_X = \infty$. Its shows there is a correlation between the μ and σ parameters and so poor estimates of σ will lead

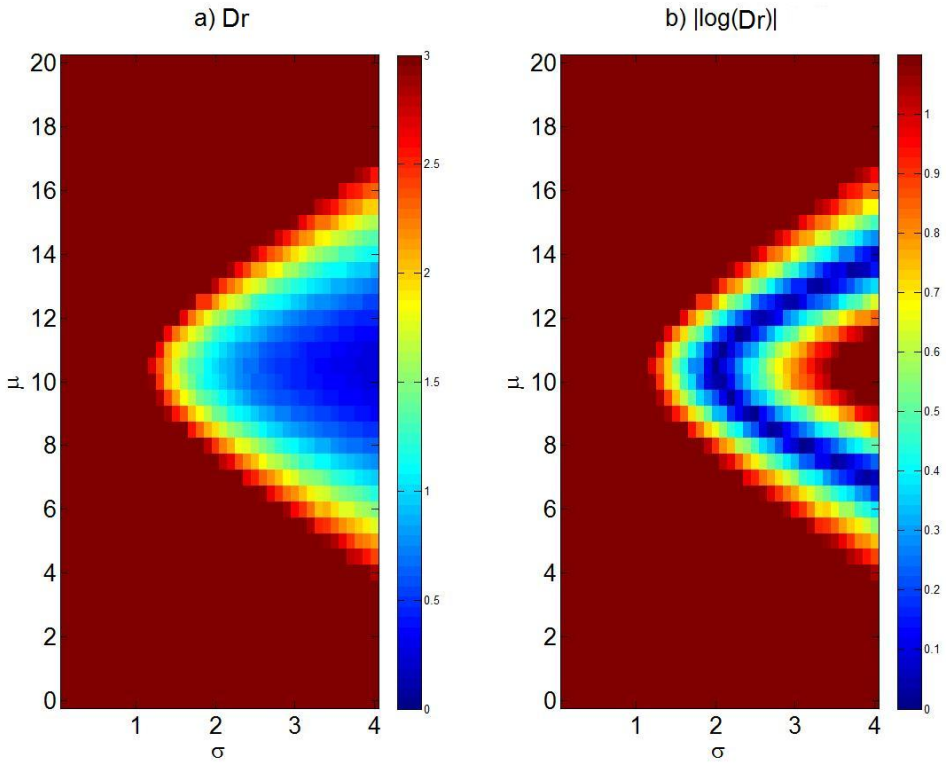

Figure 7.2: $\mathrm{Dr}(D_{[100]})$ over the range of the candidate sets μ and σ. a) shows a heat-map where the range of plausible values is everything not in the deep red. b) shows $|\log(\mathrm{Dr}(D_{[100]}))|$. This has an arrow shaped region for which the joint parameter choice suggest a good fit.

to poor estimates of μ (and vice versa). Figure 7.2b shows $|\log(\mathrm{Dr}(D_{[100]}))|$. This has an arrow shaped region for which the joint parameter choice suggest a good fit.

One alternative metric which could be used for fitting instead of the Mahalanobis distance is the Euclidean distance/ Mean Squared Error:

$$\sum_{k=1}^{100}(\mu_i - X_k)^2$$

Figure 7.3a and 7.3b shows a comparison between parameter estimation using the Mahalanobis and Euclidean distance. Unlike the Mahalanobis distance the Euclidean distance has no (co)variance component. Perhaps unsurprisingly the Euclidean distance is unable to infer the value of σ. For our case any potential distance metric comparing parameter choice is only useful if it is informative for variance learning.

7.2.4 Taking Differences

Figure 7.3a shows that the estimate for the mean, μ, influences the estimate for σ. When trying to learn about the variance parameter, we therefore take differences of observations $X'_k = X_{k+1} - X_k$ for all $k = \{1, ...n-1\}$ to remove the mean effects. We then can compute the standardised Mahalanobis distance $\mathrm{Dr}(D_{[n]})$ on the vector of differences, $D'_{[n]} = X'_1, \ldots X'_{n-1}$.

Figure 7.3c shows $|\log(\mathrm{Dr}(D_{[n]}))|$ in the case of differences of observations. Here the estimate of σ is independent of the choice of μ and so we can separate learning about means and variances. Figure 7.3d shows the Euclidean distance in the case of differenced observations. As the Euclidean distance has no covariance component and the mean effects have been removed, this shows no signal at all. Again here the Euclidean distance is ineffective in trying to learn about variance parameters

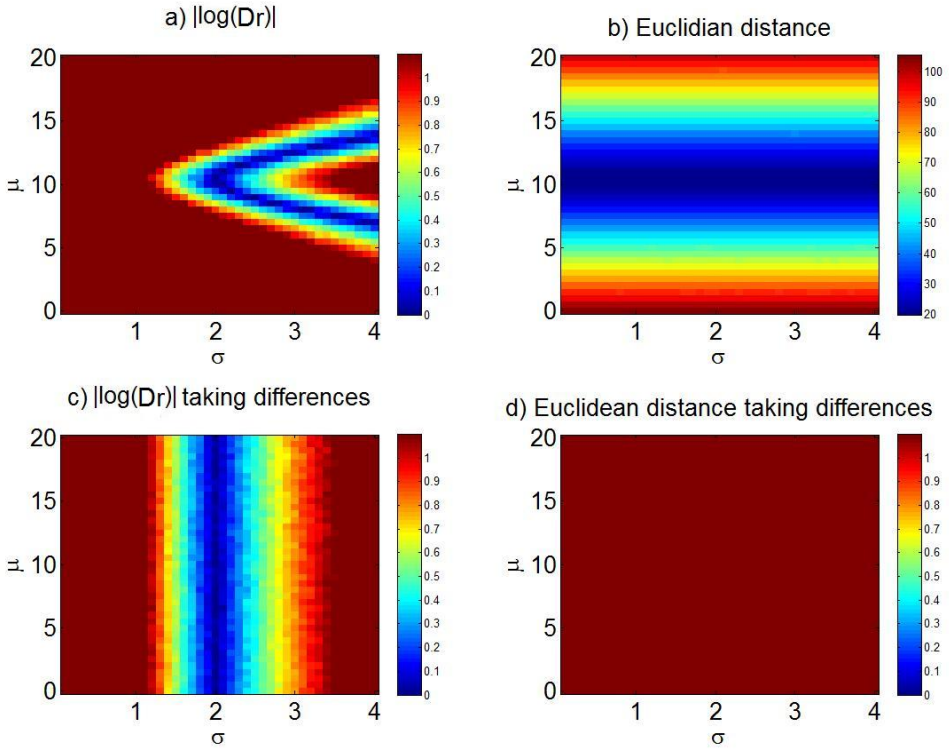

Figure 7.3: Comparison of the Mahalanobis and Euclidean distance over range of candidate values for μ and σ. a) shows a heat-map where the range of "not implausible" values is everything not in the deep red. We see that the estimate for the mean, μ it influences the estimate for σ . b) shows the same plot as a) but using the Euclidean distance. As the Euclidean distance has no covariance component this shows no signal in the variance component. c) shows $|\log(\mathrm{Dr}(D_{[n]}))|$ in the case of differences of observations. d) shows the same plot as c) but using the Euclidean distance. As the Euclidean distance has no covariance component and the mean effects have been removed, this shows no signal at all.

7.3 Mahalanobis learning: Multivariate normal with correlated observations

Instead of n independent observations we will consider observations now from an n dimensional multivariate normal:

$$X \sim N_n(\mu, \Sigma)$$

where Σ is parameterised by a constant variance σ^2 and constant covariance γ:

$$\Sigma = \begin{pmatrix} \sigma^2 & \gamma & \dots & \gamma \\ \gamma & \sigma^2 & & \vdots \\ \vdots & & \ddots & \gamma \\ \gamma & \dots & \gamma & \sigma^2 \end{pmatrix}$$

where $\gamma = \rho\sigma^2$, $0 \leq \rho \leq 1$ is the correlation. We remove the mean effects by looking at the difference of observations and compute the Mahalanobis distance as in equation 4.7.6.

Consider estimating the correlation parameter ρ. For example let the "real" $\rho = 0.5$ (with $\mu = 10$ and $\sigma = 2$ as before) and then examine $|\log(\mathrm{Dr})|$ over the candidate set:

$$\rho_X = 0, 0.05, 0.1, \dots, 0.95.$$

Treating μ and σ as known, figure 7.4a shows $|\log(\mathrm{Dr})|$ for fixed σ. The minimum of $|\log(\mathrm{Dr})|$ is $\rho \approx 0.5$. However looking jointly at σ and ρ, figure 7.4b shows a region of low $|\log(\mathrm{Dr})|$ making it very difficult to distinguish between σ and ρ.

Figure 7.5 shows profile negative log likelihood for the same data set. The plot also shows confounding making it difficult to distinguish between σ and ρ. The confounding problem in the likelihood case is also worse than the Mahalanobis fit for the same problem. For the full scale problem we do not have a well defined likelihood.

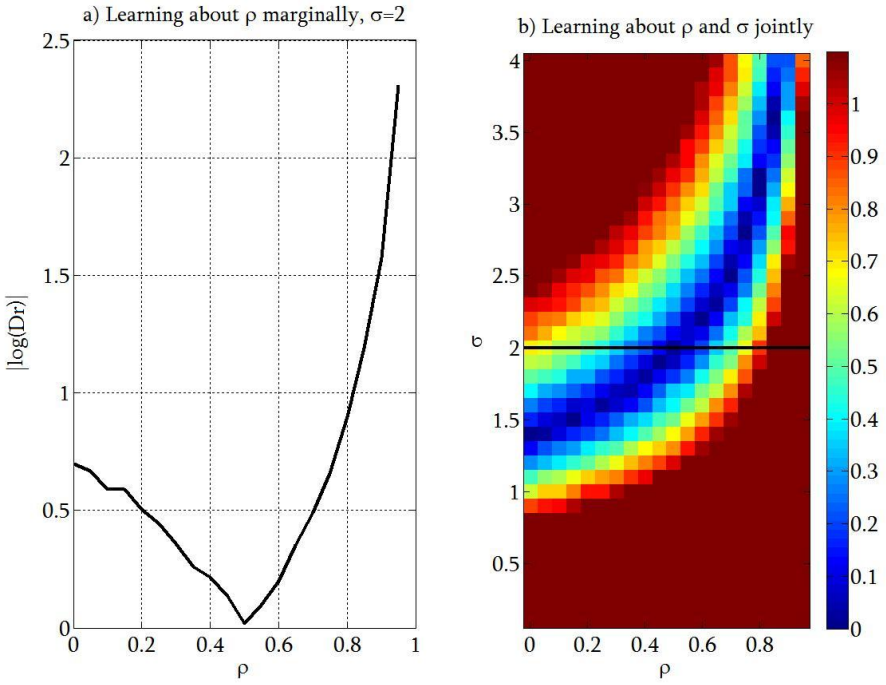

Figure 7.4: $|\log(\mathrm{Dr})|$ over range of candidate values for ρ and σ with correlated observations where the "real" $\rho = 0.5$. a) shows $|\log(\mathrm{Dr})|$ for fixed $\sigma = 2$ the minimum is at the correct value of ρ . b) shows heat-map where the range of plausible values is everything not in the deep red. This shows a region of low $|\log(\mathrm{Dr})|$ making it very difficult to distinguish between σ and ρ

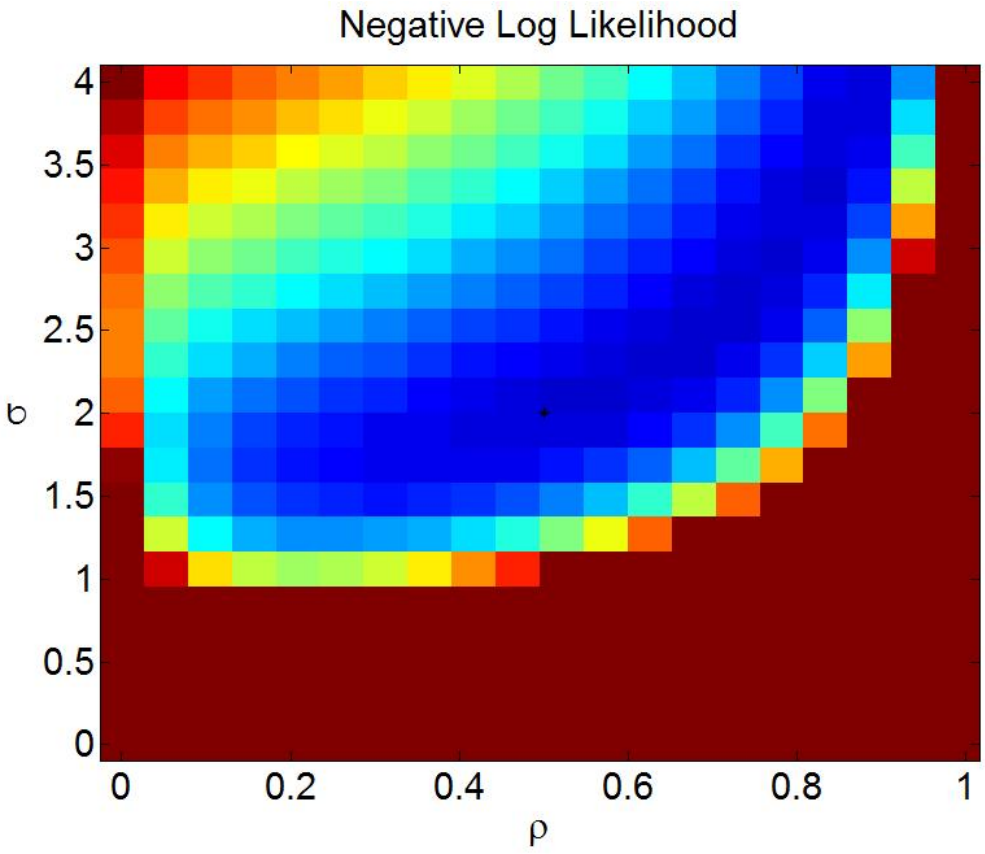

Figure 7.5: Heat-Map of the negative log likelihood over range of candidate values for ρ and σ with correlated observations where the "real" $\sigma = 2$; $\rho = 0.5$. This plot shows confounding making it difficult to distinguish between σ and ρ

7.4 Mahalanobis Variance Learning for the Local Effects Model

We now consider the more challenging case, learning about the variances within the local effects model given in section 3.2. The same approach is taken as in the example in the previous sections learning about the parameters of a multivariate normal distribution, in section 7.2.

Observational data is compared to a candidate set of parameters, Σ_r, of p, potential values for the local corrosion variances:

$$\Sigma_r = \{\Sigma_{r1}, \ldots, \Sigma_{rk}, \ldots \Sigma_{rp}\}.$$

We then wish to try to decide which of our candidate set of parameters is best the best fit with the observed data. Again the Mahalnobis distance is the criteria used to assess goodness of fit.

Unlike in the case of the normal distribution we cannot directly infer the mean and variance of the observations from the parameter choice Σ_r. Instead for each choice for the local effects variance, Σ_{rk}, we simulate realisations of the observation equation 3.3.11, treating the global effects variances as known. The simulations can be used to give estimates for $E(Y)$ and $\mathrm{Var}(Y)$. These can be used together with observations to compute the Mahalanobis distance and assess how "good" a particular choice of local variance is at explaining the data.

For each parameter choice of the local effects variance, Σ_{rk}, we simulate and assess the fit to data using the Mahalanobis distance:

$$\mathrm{Dis}(Y) = (Y - E(Y : \Sigma_{rk}))^T \mathrm{Var}(Y : \Sigma_{rk})^{\dagger} (Y - E(Y : \Sigma_{rk}))$$

where $E(Y : \Sigma_{rk})$ and $\mathrm{Var}(Y : \Sigma_{rk})$ are the simulated estimates of $E(Y)$ and $\mathrm{Var}(Y)$ respectively given a particular choice of local variance Σ_{rk}.

The range of parameter choices for which $\mathrm{Dr}(Y)$ is near to 1 (its expected value) are the best supported by the data.

7.5 Variance learning for local and global effects terms together

In the previous chapters, 5 and 6 we have shown that, for a fixed value of the local effect variance, we can update our beliefs about the global effects variance using Bayes linear variance learning. We cannot use the same approach to learn about the local effects due to the non-linearity within the observation equation. In the previous section, 7.4, a method for estimating local effects variances given fixed global effects was described.

To learn about local and global effects variances together we could specify joint candidate values over a grid as in section 7.2.3 and test each one using the Mahalanobis distance. However, this could be computationally intensive due to the need to runs simulations for each choice of parameters.

Alternatively given observational data Y, we specify a candidate set of local effects variances as in section, 7.4:

$$\Sigma_r = \{\Sigma_{r1}, \ldots, \Sigma_{rp}\}.$$

Then for each candidate choice of Σ_{rk} we use a two stage process

1. firstly we treat Σ_{rk} as known and use the value to learn about the global effect variance $E_D(\mathcal{M}(W) : \Sigma_{rk})$ using Bayes linear variance learning described in section 5.6.

2. secondly use adjusted values for the global effects variance together with our candidate local corrosion variance Σ_{rk}. Given the set of global and local variances we can generate realisations of our model. We can then generate large numbers of realisations and used them to calculate us empirical estimates of $E(Y : E_D(\mathcal{M}(W)), \Sigma_{rk})$ and $\mathrm{Var}(Y : E_D(\mathcal{M}(W)), \Sigma_{rk})$ as described in section 7.4.

We then use the Mahalanobis distance to find the candidate variance parameters

which best fits with our observational data Y:

$$\begin{aligned}\mathrm{Dr}(Y) = &(Y - E(Y : E_D(\mathcal{M}(W)), \Sigma_{rk}))^T \mathrm{Var}(Y : E_D(\mathcal{M}(W)), \Sigma_{rk})^\dagger \\ &\times (Y - E(Y : E_D(\mathcal{M}(W)), \Sigma_{rk})).\end{aligned}$$

where $E(Y : E_D(\mathcal{M}(W)), \Sigma_{rk})$ and $\mathrm{Var}(Y : E_D(\mathcal{M}(W)), \Sigma_{rk})$ are to be simulated for each of the candidate variances parameters.

Both stages, the Bayes linear updating of the global effects variances and the Mahalanobis fitting stage need to be simulated for each of the candidate set, so that we need to do $2p$ simulation runs. This could be computationally demanding for p large. However it is still preferable to joint Mahalanobis fitting of the global and local over a grid which would require $O(p^2)$, simulation runs.

7.6 Algorithm for updating mean and variance parameters in the model

We now have a methodology to update all the variance parameters and update mean parameters within our complete model described in section 3.3.4.

The procedure followed to apply the model consists of the stages as explained below. Firstly we make a prior specification and carry out simple diagnostic checks to confirm consistency of the data and priors. We then update variance and covariances in the model. Thereafter we update means and perform model diagnostics assessing model fit.

Stage 1: prior specification

First we make our prior specification for the, mean, variances, covariances and fourth order quantities in the whole model and the candidate set of local corrosion variances we wish to examine (section 4.2).

Stage 2: data diagnostics

We then check that the prior specification is consistent with the data by computing the discrepancy $\mathrm{Dis}(X)$ (section 4.7).

Stage 3: updating general corrosion variances

Given a local corrosion variance parameter, Σ_r, we can compute the Bayes Linear adjusted expectation, $E_D(\mathcal{M}(W))$, for the general corrosion variance. Given a set of local variance parameters we can then in the same way compute the adjusted expectation for each local variance. This results in giving us a candidate set variance parameters for each local variance parameter chosen (section 6.2).

Stage 4: updating local corrosion variances

We then need to decide which of the candidate variance parameters best fits the observed data. We therefore compute the Mahalanobis distance for each set of variances:

$$\begin{aligned}\mathrm{Dr}(Y) = (Y - E(Y : E_D(\mathcal{M}(W)), \Sigma_{rk}))^T \mathrm{Var}(Y : E_D(\mathcal{M}(W)), \Sigma_{rk})^\dagger \\ \times (Y - E(Y : E_D(\mathcal{M}(W)), \Sigma_{rk})).\end{aligned}$$

The set of variances which best fit the data are then chosen by using $\min(|\log(\mathrm{Dr})|)$ (section 7.5).

Stage 5: further data diagnostics

We check the adjusted expectations are plausible by computing the adjustment discrepancy on the updated variances. (section 4.7)

Stage 6: mean updating

Using our updated variance structure we can generate new simulations of the full model and use them learn about the system state (section 4.1).

Stage 7: further data diagnostics

We then carry out further data diagnostics once we have updated the means and variances to check the consistency of updated values. We can compute the adjustment discrepancy on both the BL updated of the system state and the general variances.

Stage 8: explore reliability of procedure

We could carry out one further stage to give an idea of the uncertainty of our estimates since there is no natural way of estimating the uncertainty in the Mahalanobis distance. We can generate multiple synthetic datasets using the updated parameter values. These datasets can then be used to assess the reliability of the fitting procedure.

7.7 Example: Updating the variances for offshore platform application

We continue the offshore platform example from previous chapters (previously discussed in sections 1.4.1, 3.5, 4.9, 5.7.1, 5.7.2 and 6.4).

In the previous chapters example (section 6.4) we showed an example learning about general effects variance for fixed local effects variances. We now extend this example to consider learning jointly about both local and general effects together.

In the examples we use the same 50 synthetic data sets used the example in section 6.4. We carry out each of the stages described in the algorithm above in section 7.6.

7.7.1 Synthetic Data

An application of the method to analysis of inspection data from a full-scale offshore platform is now considered. In section 6.4 we showed an example using synthetic data of Bayes linear learning about the general corrosion variances, treating the local variance parameters as known. Here we consider the same example but learn about both local and general variance parameters.

Prior Specification

Stage one in the algorithm is the prior specification. The prior mean specification (corrosion rates and wall thicknesses) is the same as used in for the offshore platform

example in section 4.9. Our prior variance specification is given as:

$$\Sigma_Y = 0.167 \qquad \Sigma_X = 0.1 \qquad \Sigma_\alpha = 0.001$$

We specify a prior candidate set for the local corrosion variance:

$$\Sigma_r = 0, 0.01, \ldots 0.1, 0.2 \ldots 1 \tag{7.7.1}$$

The simulated real system has variances:

$$\Sigma_Y = 0.167 \qquad \Sigma_X = 0.1 \qquad \Sigma_r = 0.1 \qquad \Sigma_\alpha = 0.001$$

As in the example in the previous chapter, section 5.7.1 we will investigate how well we can recover the "real" system variances.

Stage two in the algorithm is the data diagnostics. Figure 6.1 (in the previous chapter section 6.4) shows the point wise prior discrepancy of one realisation of synthetic data. Stage three updating the general corrosion variances was also discussed in the previous chapter.

Variance Learning

Stage four in the algorithm is the joint updating local corrosion variances given general corrosion variances. In section 6.4.2 in the previous chapter we examined explored learning about the general corrosion variances using 50 synthetic datasets. We now illustrate joint Bayes linear and Mahalanobis variance learning for the simulated example in figures 7.6 and 7.7 using the same datasets.

Given a set of location corrosion variances we can compute the Bayes Linear adjusted expectation for the general corrosion variances. Figure 7.6 shows Bayes linear variance learning for wall thickness variance, Σ_{W_X}, as a function of the candidate set of local corrosion variances specified in equation 7.7.1. The simulated example has true values of Σ_{W_X} and Σ_{V_r} both 0.1. We can see that for the true local corrosion parameter we get do indeed predict the correct general corrosion variance.

After the computing the adjusted expectation for the general corrosion variances for each local corrosion variance in the candidate set we can then ask which of these combinations is best. We generate simulations of the corrosion model (section 3.5)

and compute empirical expectations and covariance matrices used to compute the Mahalanobis distance for each of the candidate set of variances. Figure 7.7 shows the standardised Mahalanobis distance, $\mathrm{Dr}(Y)$, as a function of local corrosion variance, Σ_{V_r} for the same example. The minimum value of $\log(\mathrm{Dr}(Y))$, gives parameter estimates of gives parameter estimates of $\Sigma_{V_r} = 0.09$ and from figure 7.6 $\Sigma_{W_X} = 0.11$.

We have therefore managed to get a good estimate for both the local and general corrosion variances jointly. We can then proceed to stages 5-7 in the algorithm, mean updating and further data diagnostics. These stages have been examined in previous chapters, section 4.9.

Prior Sensitivity to Σ_X

We can examine the sensitivity of the method to the prior specification. Figure 7.8 shows variance learning for the general and local corrosion variances, given 3 prior choices for $\Sigma_X = \{0.05, 0.1, 0.2\}$. We see the range of updating values for the local corrosion vary from $\Sigma_{V_r} = [0, 0.3]$ and the corresponding general variance range from $\Sigma_{W_X} = [0.05, 0.16]$.

The adjusted expectation of the general corrosion variance in all cases has moved from the prior towards the true value. However we do not have enough data to disregard the prior and use the data mean. The consequence of this is under prediction of the general corrosion variance in the case of prior which is "too small" and similar variances which are "too big" in opposing situation. Choice of local corrosion variance in these situations seems to want to try to match the total variation in the model. When the general corrosion variance is too high the location corrosion variance is small and vice versa. The effect of this is shifting whether the variability is modelled into the local effects model or the general effects model.

7.8 Example: Historical Inspection Data

We now consider an example of learning about the general and local corrosion variance parameter in the case of analysis the real historical inspection data described in section 1.4.1. A system of four corrosion circuits is modelled, consisting of a total

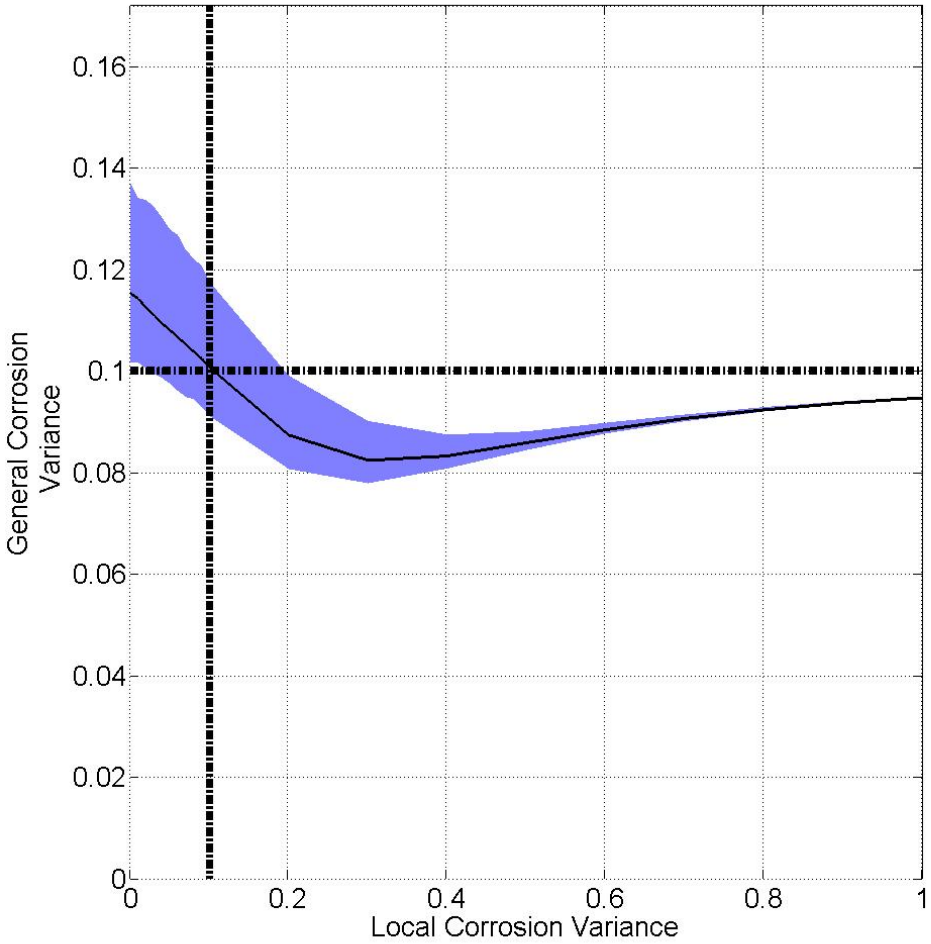

Figure 7.6: Bayes linear variance learning for wall thickness variance, Σ_{W_X}, as a function of local corrosion variance, Σ_{V_r}. The true values of Σ_{W_X} and Σ_{V_r} are both 0.1, as shown by the dashed horizontal and vertical lines. The mean estimate for Σ_{W_X} is shown as a solid line, and the blue shaded region corresponds to a 90% uncertainty band for Σ_{W_X} bounded by the 5th and 95th percentiles derived from simulation using 50 sets of synthetic data.

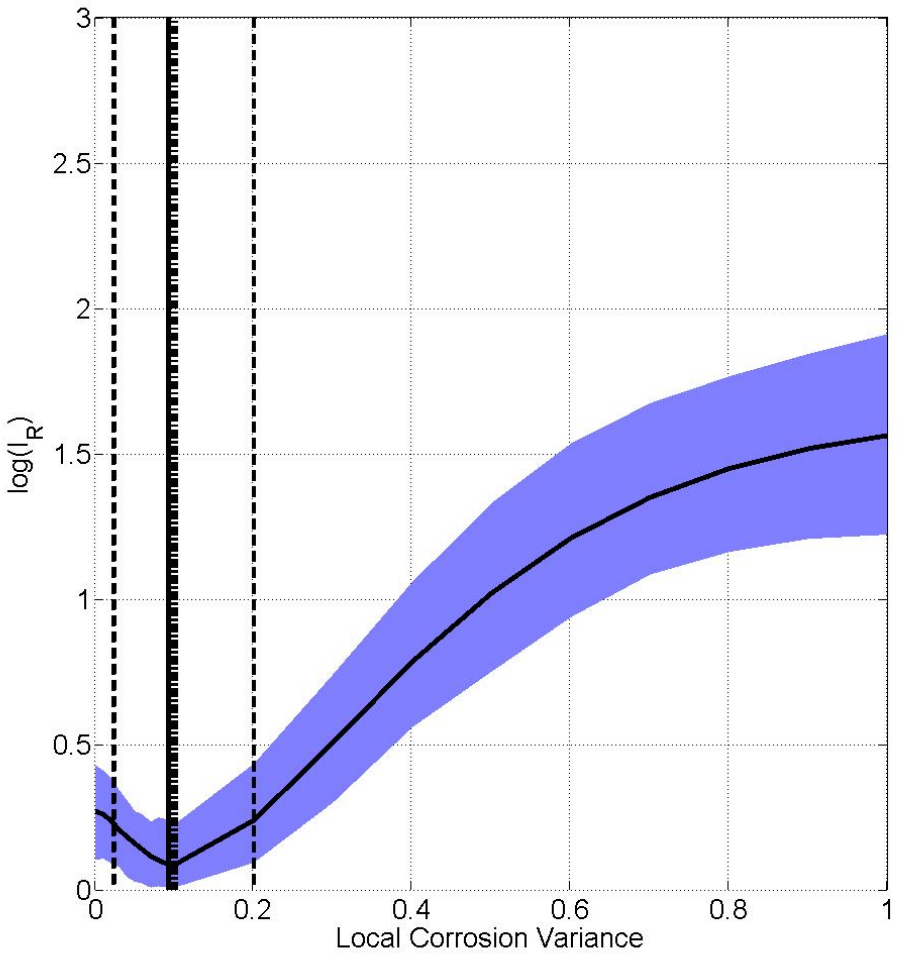

Figure 7.7: The standardised Mahalanobis distance, $\mathrm{Dr}(Y)$, as a function of local corrosion variance, Σ_{V_r}. The mean estimate for $\mathrm{Dr}(Y)$ is shown as a solid line, and the blue shaded region corresponds to a 90% uncertainty band for $\mathrm{Dr}(Y)$ bounded by the 5th and 95th percentiles derived from simulation using 50 sets of synthetic data. The minimum value of $\mathrm{Dr}(Y)$, suggests $\Sigma_{V_r} = 0.09$ and corresponding $\Sigma_{W_X} = 0.11$, (from figure 7.6). Vertical lines indicate the mean (solid) and 5th and 95th percentiles (dashed) for the particular choice of Σ_{V_r} in individual realisations. The true values of Σ_{W_X} and Σ_{V_r} are both 0.1.

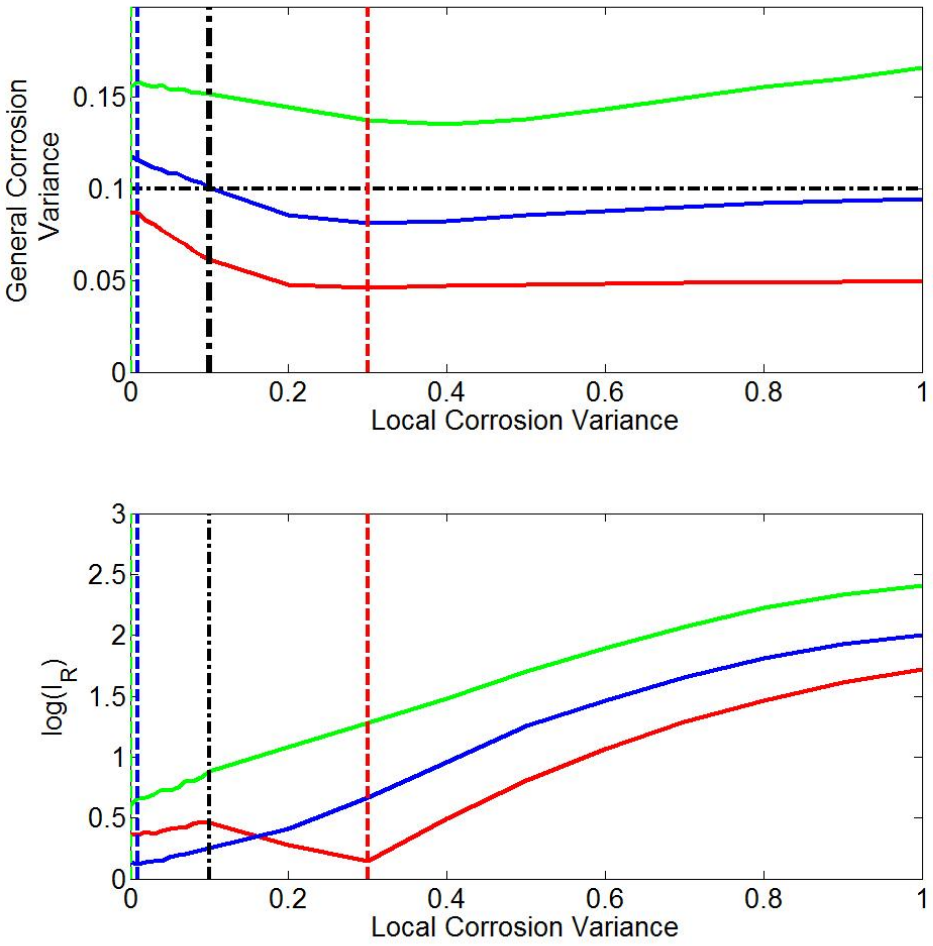

Figure 7.8: Variance learning for the synthetic data with 3 different prior choices. The red line shows joint variance learning for $\Sigma_{W_X} = 0.05$, the blue line shows $\Sigma_{W_X} = 0.1$ and the green line shows $\Sigma_{W_X} = 0.2$. The coloured vertical lines indicate the local corrosion variance estimate for each prior choice. The true values of Σ_{W_X} and Σ_{V_r} are both 0.1, as shown by the black dot dashed horizontal and vertical lines

of 64 pipe-work weld components, data is given in appendix A. For problems of this size simulation and modelling can take around an hour on a standard laptop. For bigger practical sized problems such as the modelling of an entire oil platform the scale of simulations would necessitate access to more powerful computing and multi/core computing. The nature of the simulation procedure means that this approach is extremely easy to parallelise and send to multiple cores.

Prior specification is the same as given in previous section 7.7.1, for the synthetic data. Figure 7.9 shows the point wise discrepancy of the historical inspection data. At time 12 we see large discrepancies suggesting there is either something wrong with our prior specification at these points or maybe different measurement technique. Using data spread over many years means that consistent data gathering is problematic, instruments, operators and procedures are all subject to change. In section 4.9 we discussed updating the expected wall thickness and corrosion rates for these data. We now consider learning about the general and local corrosion variances.

We assume that we can specify the measurement error variance directly from our knowledge of the inspection processes and equipment. Even when instruments have changed, the device used is recorded. We therefore wish to estimate the general and local corrosion variances. The procedure is the same as discussed in the previous section 7.7.1. First a set of candidate local variances decided. The adjusted expectation for the general corrosion variances is computed for each of the local variances. Each candidate combination of local and general corrosion variances can then be compared against data using the Mahalanobis distance.

The updated general corrosion variances Σ_{W_X} and local variances Σ_{V_r} results are shown in Figure 7.10 (to be compared with 7.6 and 7.7). The fitting procedure suggests that a local corrosion variance of 0.3 and general corrosion variance of 0.1 are best. Unfortunately we have no estimate of the uncertainty around this estimate. An estimate of the uncertainty around the combination of local and general corrosion variances can be found by computing the adjusted variance, however there is no easy way to compute the uncertainty around the Mahalanobis distance. Sensitivity to the prior specification can be explored in the same way as in section 7.7.1.

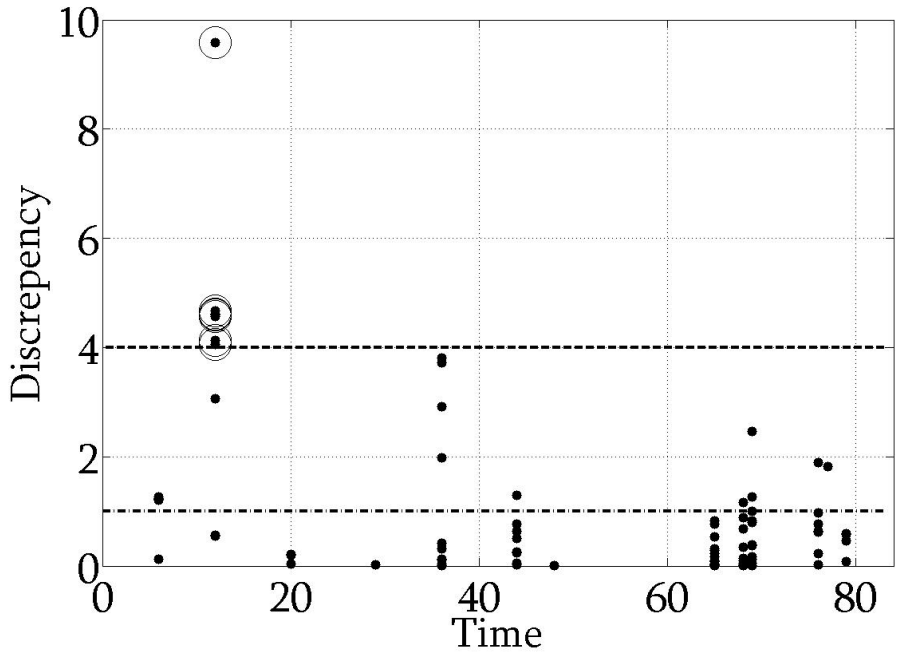

Figure 7.9: The point wise discrepancy of one realisation of the real historical inspection data. The expected value of $\mathrm{Dis}(Y_{ct})$ is 1, shown as a horizontal line. Also shown is the horizontal line corresponding to $|1 - \mathrm{Dis}(Y_{ct})| = 3\sqrt{2}$ the $3\sigma_{\mathrm{Dr}Y}$ rule under normality, serving as a warning limit for unusually large values of discrepancy. At time 12 we see large discrepancies suggesting there something wrong with our prior specification at these points.

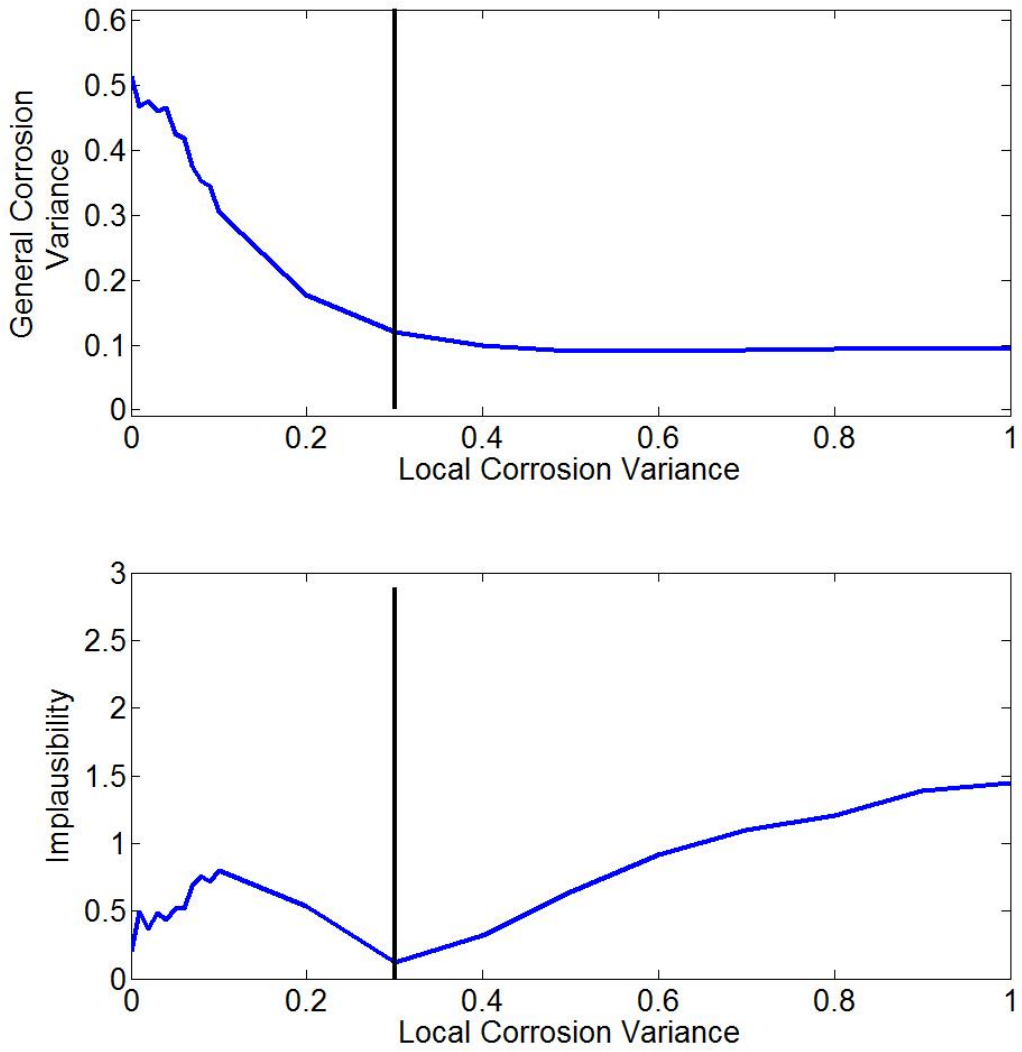

Figure 7.10: Bayes linear variance learning for the historical inspection data. The top plot shows the wall thickness variance, Σ_{W_X}, as a function of local corrosion variance. The bottom plot shows the standardised Mahalanobis Distance, Dr, as a function of local corrosion variance. The vertical line shows the minimum log(Dr) of suggesting a local corrosion variance of 0.3 and corresponding general corrosion variance of 0.12.

One approach to quantifying this uncertainty in the Mahalanobis distance could be done by using the updated model estimates to generate new synthetic data sets and refit the model. This approach would give an idea of the reliability of the results.

7.9 Covariance Learning

If we had enough data we could attempt to learn about the correlations/covariance parameters within the model as well as means and variances. One could imagine a similar approach to that of variance learning using exchangeability.

7.9.1 Exchangeability in Time

First consider the case of full inspections where we have long time series data. We make the assumption as for the variance learning case that the residuals ,ϵ_{Xct}, are exchangeable in time, giving a representation statements for each component:

$$\epsilon_{Xct} = \mathcal{R}_t(\epsilon_{Xc})$$

We also make the assumption that the product of the residuals for a pair of components, c and c', are exchangeable in time thus, $c \neq c'$:

$$\mathcal{R}_t(\epsilon_{Xc})\mathcal{R}_t(\epsilon_{Xc'}) = \mathcal{C}_{Xt} = \mathcal{M}(\mathcal{C}_X) + \mathcal{R}_t(\mathcal{C}_X)$$

where we make the prior specification:

$$E(\mathcal{M}(\mathcal{C}_X)) = \Sigma_{Xcc'} \qquad \text{Var}(\mathcal{M}(\mathcal{C}_X)) = \Xi_{\mathcal{C}_X} \qquad \text{Var}(\mathcal{R}_t(\mathcal{C}_X)) = \Phi_{\mathcal{C}_X} - \Xi_{\mathcal{C}_X}.$$

7.9.2 Adjusting Beliefs in the case of equally spaced observations

We can then compute the adjusted expectation $E_D(\mathcal{M}(\mathcal{C}_X))$ using a product of differences of observations:

$$D_{cc't} = (Y^{(2)}_{ct} - 2Y^{(1)}_{ct})(Y^{(2)}_{c't} - 2Y^{(1)}_{c't}),$$

then $D_{cc'}$:

$$D_{cc'} = \begin{pmatrix} D_{cc'1} \\ \vdots \\ D_{cc'T} \end{pmatrix}$$

In this way we can update our beliefs about individual covariances within the model by computing the adjusted expectation, $E_{D_{cc'}}$:

$$E_{D_{cc'}}(\mathcal{M}(\mathcal{C}_X)) = E(\mathcal{M}(\mathcal{C}_X)) + \mathrm{Cov}(\mathcal{M}(\mathcal{C}_X), D_{cc'})\mathrm{Var}(D_{cc'})^{-1}\left(D_{cc'} - E(D_{cc'})\right)$$

However, in practice we have irregularly spaced data with short time series. In order to learn about a covariance we need sets of 3 observations on both components. Therefore in this case it is highly unlikely we have enough data between pairs of components to be able to learn about correlation. If we attempted to use exchangeability assumptions across components we have the problem that different components are not observed at the same time. If we assume some form of structure to the correlation matrix such as a distance then we may be able to imagine exchangeability between equidistant components. However, for a large covariance matrix this would still become infeasibly complicated, even if you had enough data to make inferences and could justify the assumptions.

7.9.3 Mahalanobis covariance learning

An alternative approach would be to use Mahalanobis fitting as we did in section 7.4. We choose a candidate set, Σ_X of correlation matrices and assess which fits the data best using the Mahalanobis distance:

$$\Sigma_X = \{\Sigma_{X1}, \Sigma_{X2}, \ldots, \Sigma_{Xp}\}$$

We partition the data D into 2 parts, $D_{\{c_1\}}$, and $D_{\{c_2\}}$ which contain observations from disjoint sets of components $\{c_1\}$, $\{c_2\}$.

Each candidate choice of Σ_{Xi} is treated as a prior specification. We can then generate a large number realisations of the model and calculate empirical estimates for all relevant means, variances and covariances of interest. We then use the data $D_{\{c_1\}}$ from the set of components $\{c_1\}$ to learn about the set $\{c_2\}$ by computing

Bayes linear adjustment, $E_{D_{\{c_1\}}}(Y_{\{c_2\}})$:

$$\begin{aligned} E_{D_{\{c_1\}}}(Y_{\{c_2\}}) &= E(Y_{\{c_2\}}) + \text{Cov}(Y_{\{c_2\}}, D_{\{c_1\}})[\text{Var}(D_{\{c_1\}})]^{-1}(D_{\{c_1\}} - E(D_{\{c_1\}})) \\ \text{Var}_{D_{\{c_1\}}}(Y_{\{c_2\}}) &= \text{Var}(Y_{\{c_2\}}) - \text{Cov}(Y_{\{c_2\}}, D_{\{c_1\}})[\text{Var}(D_{\{c_1\}})]^{-1}\text{Cov}(D_{\{c_1\}}, Y_{\{c_2\}}) \end{aligned}$$

Since the data is partitioned into 2 disjoint sets of components the only learning which occurs is directly related to the strength of correlation within the system. This approach is similar to a cross validation.

Adjusted estimates for the components $\{c_2\}$ can be compared against the real observations using the Mahalanobis distance. We test the how well each choice of correlation matrix fits the data $D_{\{c_2\}}$, from equation 4.7.7:

$$\text{Dr}(D_{\{}c_2\}) = (D_{\{c_2\}} - E_{D_{\{c_1\}}}(Y_{\{c_2\}}))^T[\text{Var}_{D_{\{c_1\}}}(Y_{\{c_2\}})]^{-1}(D_{\{c_2\}} - E_{D_{\{c_1\}}}(Y_{\{c_2\}}))$$

We could test each choices within the candidate set of correlation matrices and find which best fits with the observed data. The choice on how to partition the data into these 2 disjoint sets will have an effect on how the learning occurs. An approach of repeated resampling of these 2 sets will give an estimate of the sensitivity of the data to this choice and the associated uncertainty we should attached to estimates.

7.9.4 Example: Covariance learning Offshore Platform example

Using the synthetic example continuing on from previous chapters (previously discussed in sections 1.4.1, 3.5, 4.9, 5.7.1, 5.7.2, 6.4 and 7.7).

7.9.5 Synthetic Data

We use the same 50 synthetic data sets used the example in section 6.4. The prior mean (corrosion rates and wall thicknesses) specification is the same as used in for the offshore platform example in section 4.9. The prior variance specification is the same as used in the offshore platform example in section 7.7.

Candidate correlation set

We generate the correlation matrices using the approach discussed in section 4.5.3, from equation 4.5.4:

$$\Sigma_{\Theta cc'} = \sigma_{\Theta c}\sigma_{\Theta c'}\left(\rho_0 + (1-\rho_0)\rho_{\text{Cir}}\delta_{\text{CirCir}'} + (1-\rho_0-(1-\rho_0)\rho_{\text{Cir}})e^{-\nu s_{cc'}}\right)$$

We specify a candidate set of ρ_{Cir}:

$$\rho_{\text{Cir}} = [0.001, 0.01, 0.05, 0.1, 0.2, \ldots, 0.7]$$

with fixed:

$$\rho_0 = 0.2 \qquad\qquad \nu = 0.05$$

The real system has parameter values:

$$\rho_0 = 0.2 \qquad \nu = 0.05 \qquad \rho_{\text{Cir}} = 0.1$$

Figure 7.11 shows Mahalanobis distance, Dr, as a function of the circuit covariance parameter, ρ_{Cir} for the simulated example. The expected value of Dr, shown as a horizontal line is 1, suggesting a value of $\rho_{\text{Cir}} = 0.1$. The true value of ρ_{Cir} is also 0.1 so on average we pick the correct value. Even in this cases with synethitic data we see that we could have chosen a correlation anywhere from $[0 - 0.4]$.

7.9.6 Historical Inspection Data

Consider an example of learning about the general and local corrosion variance parameters in the case of analysis of the real historical inspection data described in section 1.4.1. A system of four corrosion circuits is modelled, consisting of a total of 64 pipe-work weld components, data is given in appendix A. Figure 7.12 shows the Mahalanobis fitting for the covariance structure varying over the circuit covariance parameter, ρ_C. Bounds are generated by resampling and fitting with different sets of data and fitting sets. The bounds are massive and Dr is large suggesting there is insufficient data or the model fit is too poor to be able to chose a correlation parameter. It is possible that with more data that there we could get an estimate for the correlation parameter. However we are reaching the limit of insights from the data.

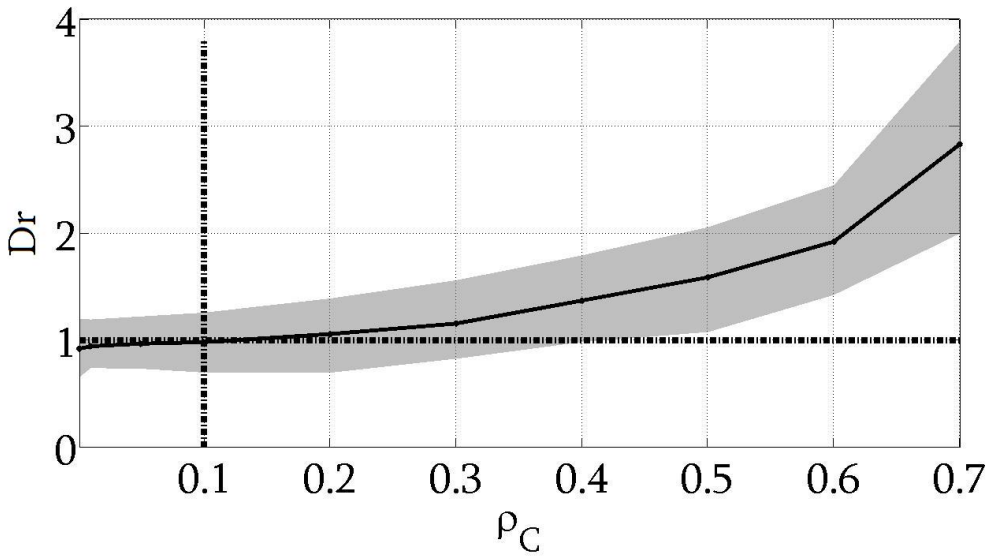

Figure 7.11: Mahalanobis distance, Dr, as a function of the circuit covariance parameter, ρ_C. The mean estimate for Dr is shown as a solid line, and the shaded region corresponds to a 90% uncertainty band for Dr bounded by the 5th and 95th percentiles derived from simulation. The expected value of Dr, shown as a horizontal line is 1, suggesting a value of $\rho_C = 0.1$. The true value of ρ_C is also 0.1.

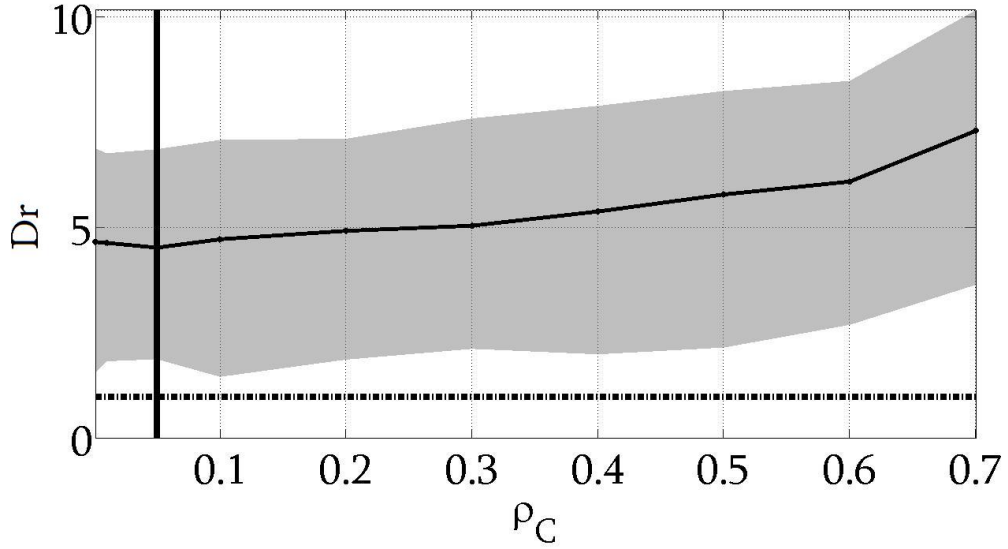

Figure 7.12: Mahalanobis fitting for the covariance structure varying over the circuit covariance parameter, ρ_C. The bounds are quite wide and not centered around 1 suggesting model fit is far for perfect. However a choice of $\rho_C = 0.05$ would seem to be best

Chapter 8

Efficient Inspection

We now have a model which can be used to predict the current state of the system at any time. We also have a way of updating our beliefs about mean and variance structures using Bayes linear adjustments. We will now consider the problem of designing an efficient inspection scheme. Its value is assessed in terms of reducing uncertainty about the system state, thus minimising potential losses from component failure. The following is discussed in [Randell et al., 2010] and [G.Hardman, 2007].

We begin by formalising the decision problem, section 8.1. We then discuss the utility/loss criteria for a given design, section 8.2, Before showing how to evaluate the expected loss, in section 8.3 and evaluate expected under certain normality assumptions, in section 8.4. Once we have a method of evaluating the worth of a particular design we then discuss the problem of efficient searching of the design space, section 8.5. We then extend the design selection in section 8.6 to consider a design approach also incorporating variance learning. We then consider an example applied to the offshore platform data in section 8.7. Work described in sections 8.1 to 8.4 are based on work described in [G.Hardman, 2007], which is then extended in the subsequent sections to incorporate variance learning.

8.1 The decision problem

We first consider the value of inspecting a single component using a utility based criterion. For collections of components, summation over components allows evalu-

ation of complete designs.

To simplify the inspection problem, suppose that there are two possible outcomes, $o \in O$, namely failure, F, or survival, $\bar{F}$ per system component. System maintenance involves replacing a component, decision R, or leaving it alone, $\bar{R}$. Replacing a component incurs cost, L_R, whereas component failure costs L_F. Furthermore when a component fails it also needs replacing, i.e. $L_R \leq L_F$. This cost structure can be summarised as:

	F	$\bar{F}$
R	L_R	L_R
$\bar{R}$	L_F	0

That is, the four possible outcomes are:

1. *Replace* component when it *would have failed*; cost L_R
2. *Replace* component when it *would not have failed*; cost L_R
3. *Don't replace* component and it *fails*; cost L_F
4. *Don't replace* component and it *does not fail*; cost 0

We also need to consider that carrying out the inspection itself will incur costs. We will model these inspection costs as follows:

$$L_{IC} = L_{SUC} + n_d L_C, \tag{8.1.1}$$

where L_{SUC} is the cost of setting up the inspection, L_C is the cost of inspecting a particular component, n_d is the number of components in the inspection design d being considered. This means that as the scale of the inspection increases the cost of the inspection increases as well. These costs are added up to give the total inspection cost, L_{IC} for a particular design.

The loss incurred through inspection is independent of the other loss and so we can combine these costs with our expected loss criteria for maintenance simply by adding the costs together. The real benefit of including a measure of the inspection costs is that it gives a way of comparing designs which inspect different numbers of components so we can work out whether it is beneficial to inspection more or less components.

8.2 Utility

Utility quantifies preferences concerning different uncertain rewards on scale for which expected utility equals actual utility. Loss is negative utility. In a space of possible decisions $\Delta = \{R, \bar{R}\}$, the best decision procedure, δ^*, has maximum utility or minimum loss. For design d, yielding inspection data, Y_d, the expected loss of decision $\delta(Y_d)$ is:

$$\begin{aligned} E[L(O,\delta(Y_d))] &= E\{E[L(O,\delta(Y_d))]|Y_d\} \\ &= E\{L(F,\delta(Y_d))P(F|Y_d) + L(\bar{F},\delta(Y_d))P(\bar{F}|Y_d)\}. \end{aligned} \tag{8.2.2}$$

The component is replaced, decision R, if $E[L(O,R)|Y_d] < E[L(O,\bar{R})|Y_d]$:

$$\begin{aligned} E[L(O,R)|Y_d] &= E\{L(F,R)P(F|Y_d) + L(\bar{F},R)P(\bar{F}|Y_d)\} = L_R \\ E[L(O,\bar{R})|Y_d] &= E\{L(F,\bar{R})P(F|Y_d) + L(\bar{F},\bar{R})P(\bar{F}|Y_d)\} = L_F P(F|Y_d). \end{aligned}$$

Hence, the component is replaced if $L_R < L_F P(F|Y_d)$, i.e.:

$$\delta^*(Y_d) = \begin{cases} R \text{ if } p(F|Y_d) \geq \eta \\ \bar{R} \text{ if } p(F|Y_d) < \eta \end{cases} \qquad \text{where} \qquad \eta = \frac{L_R}{L_F}.$$

Let $q(Y_d) = P(F|Y_d)$, the probability of failure given current system observations. From equation 8.2.2 the expected loss of the optimal decision, $\delta^*(Y_d)$, is:

$$\begin{aligned} E[L(O,\delta^*(Y_d))] =& E\{L(F,\delta^*(Y_d))P(F|Y_d) + L(\bar{F},\delta^*(Y_d))P(\bar{F}|Y_d)\} \\ =& E\{L(F,\delta^*(Y_d))q(Y_d) + L(\bar{F},\delta^*(Y_d))(1-q(Y_d))\} \\ =& E\Big[L(F,\delta^*(Y_d))q(Y_d) + \\ & L(\bar{F},\delta^*(Y_d))(1-q(Y_d))\Big|q(Y_d) \geq \eta\Big] P(q(Y_d) \geq \eta) + \\ & + E\Big[L(F,\delta^*(Y_d))q(Y_d) + \\ & L(\bar{F},\delta^*(Y_d))(1-q(Y_d))\Big|q(Y_d) < \eta\Big] P(q(Y_d) < \eta) \\ =& L_R P(q(Y_d) \geq \eta) + L_F E[q(Y_d)|q(Y_d) < \eta] P(q(Y_d) < \eta) \\ =& L_R \int_\eta^1 p(q(Y_d))dq(Y_d) + L_F \int_0^\eta q(Y_d)p(q(Y_d))dq(Y_d) \\ =& L_R I_1 + L_F I_2. \end{aligned} \tag{8.2.3}$$

Therefore calculation of the expected loss of decision δ^*, requires evaluation of integrals I_1 and I_2 as explained in section 8.3 and in a specific case in section 8.4.

8.3 Evaluating expected loss

A component is deemed to have failed if the system level falls below some critical value W_C. The probability of component failure before some future time $t+k$ is:

$$q(Y_d) = P(F|Y_d) = P(X_{t+k} < W_C|Y_d),$$

where X_{t+k} is the unknown future system level at time $t+k$. To evaluate integrals I_1 and I_2 from equation 8.2.3 expressions for $q(Y_d)$ and its probability distribution, $p(q(Y_d))$ are required, which can be evaluated for any proposed design. This is achieved using a combination of Bayes linear analysis and appropriate distributional assumptions.

For inspection data Y_d, the adjusted mean and variance are:

$$\begin{aligned} E_{Y_d}(X_{t+k}) &= E(X_{t+k}) + \mathrm{Cov}(X_{t+k}, Y_d)\mathrm{Var}(Y_d)^{-1}(Y_d - E(Y_d)) \\ \mathrm{Var}_{Y_d}(X_{t+k}) &= \mathrm{Var}(X_{t+k}) - \mathrm{Cov}(X_{t+k}, Y_d)\mathrm{Var}(Y_d)^{-1}\mathrm{Cov}(Y_d, X_{t+k}). \end{aligned}$$

Note that the adjusted variance, $\mathrm{Var}_{Y_d}(X_{t+k})$ depends only on prior beliefs and the specific design, d. It does not depend on the observed inspection data, Y_d. However, the adjusted expectation, $E_{Y_d}(X_{t+k})$, depends directly on Y_d. Bayes Linear analysis is therefore also used to update beliefs about its mean, $E(E_{Y_d}(X_{t+k}))$, and variance, $\mathrm{Var}(E_{Y_d}(X_{t+k}))$. For the adjusted mean:

$$\begin{aligned} E(E_{Y_d}(X_{t+k})) =& E(E(X_{t+k}) + \mathrm{Cov}(X_{t+k}, Y_d)\mathrm{Var}(Y_d)^{-1}(Y_d - E(Y_d))) \\ =& E(X_{t+k}) + \mathrm{Cov}(X_{t+k}, Y_d)\mathrm{Var}(Y_d)^{-1}(E(Y_d) - E(Y_d)) \\ =& E(X_{t+k}). \end{aligned} \tag{8.3.4}$$

To find the adjusted variance $\mathrm{Var}(E_{Y_d}(X_{t+k}))$, use:

$$\begin{aligned} \mathrm{Var}(X_{t+k}) &= \mathrm{Var}\big(E_{Y_d}(X_{t+k}) + X_{t+k} - E_{Y_d}(X_{t+k})\big) \\ &= \mathrm{Var}(E_{Y_d}(X_{t+k})) + \mathrm{Var}(X_{t+k} - E_{Y_d}(X_{t+k})) \\ &= \mathrm{Var}(E_{Y_d}(X_{t+k})) + \mathrm{Var}_{Y_d}(X_{t+k}), \end{aligned}$$

so that:

$$\text{Var}(E_{Y_d}(X_{t+k})) = \text{Var}(X_{t+k}) - \text{Var}_{Y_d}(X_{t+k}) \tag{8.3.5}$$

These expressions for the first and second moments of X_{t+k}, and the adjusted expectation $E_{Y_d}(X_{t+k})$ given in this section are general and make no assumptions on distributional forms. In practice through searching for optimal designs requires fast calculations of I_1 and I_2. Under normal distribution assumptions (given in section 8.4), calculations can be simplified allowing fast evaluation of the expected loss.

8.4 Evaluating expected loss under normality

We now consider the evaluation of the expected loss and the integrals I_1 and I_2 from equation 8.2.3 in the case of normal distributional assumptions. Expressions for the first and second moments of X_{t+k}, and its adjusted expectation $E_{Y_d}(X_{t+k})$ are given in section 8.3. Henceforth, these quantities are assumed to be normally distributed:

$$\begin{aligned} X_{t+k}(Y_d) &\sim N(E_{Y_d}(X_{t+k}), \text{Var}_{Y_d}(X_{t+k})) \\ E_{Y_d}(X_{t+k}) &\sim N(E(X_{t+k}), \text{Var}(X_{t+k}) - \text{Var}_{Y_d}(X_{t+k})) \end{aligned} \tag{8.4.6}$$

The adjusted variance $\text{Var}_{Y_d}(X_{t+k})$ does not depend on particular realisations of data and thus can estimated from the simulated data. From equation 8.2.3 expected loss for a given design, d, is given by:

$$\begin{aligned} E[L(O, \delta^*(Y_d))] &= L_R \int_\eta^1 p(q(Y_d))dq(Y_d) + L_F \int_0^\eta q(Y_d)p(q(Y_d))dq(Y_d) \\ &= L_R I_1 + L_F I_2 \end{aligned}$$

8.4.1 Evaluating I_1

The probability of component failure is given by:

$$q(Y_d) = P(F|Y_d) = P(X_{t+k} < W_C|Y_d)$$

Therefore using the normality and standardising:

$$\begin{aligned} P(X_{t+k} < W_C|Y_d) &= P\left(\frac{X_{t+k} - E_{Y_d}(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})} < \frac{W_C - E_{Y_d}(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})}\middle| Y_d\right) \\ q(Y_d) &= \Phi\left(\frac{W_C - E_{Y_d}(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})}\right) \end{aligned} \tag{8.4.7}$$

$$\text{Let} \qquad z = \frac{W_C - E_{Y_d}(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})} \tag{8.4.8}$$

Then using equation 8.4.6:

$$\begin{aligned} \mu_z = E(z) &= E\left(\frac{W_C - E_{Y_d}(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})}\right) \\ &= \frac{W_C - E_{Y_d}(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})} \\ &= \frac{W_C - E(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})} \end{aligned} \tag{8.4.9}$$

$$\begin{aligned} \sigma_z^2 = \text{Var}(z) &= \text{Var}\left(\frac{W_C - E_{Y_d}(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})}\right) \\ &= \frac{\text{Var}(E_{Y_d}(X_{t+k}))}{\text{Var}_{Y_d}(X_{t+k})} \\ &= \frac{\text{Var}(X_{t+k}) - \text{Var}_{Y_d}(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})} \end{aligned} \tag{8.4.10}$$

So to calculate I_1:

$$I_1 = \int_\eta^1 p(q(Y_d))dq(Y_d) = P(q(Y_d) \geq \eta)$$

Then from equations 8.4.7, 8.4.8, 8.4.9 and 8.4.10:

$$\begin{aligned} P(q(Y_d) \geq \eta) &= P\left[\Phi\left(\frac{W_C - E_{Y_d}(X_{t+k})}{\text{Var}_{Y_d}(X_{t+k})}\right) \geq \eta\right] \\ &= P(\Phi(z) \geq \eta) \\ &= P(z \geq \Phi^{-1}(\eta)) \\ &= P\left(\frac{z - \mu_z}{\sigma_z} \geq \frac{\Phi^{-1}(\eta) - \mu_z}{\sigma_z}\right) \\ &= 1 - \Phi\left(\frac{\Phi^{-1}(\eta) - \mu_z}{\sigma_z}\right) \\ &= \Phi\left(\frac{\mu_z - \Phi^{-1}(\eta)}{\sigma_z}\right) \end{aligned}$$

8.4.2 Evaluating I_2

Continuing from equations 8.4.7 and 8.4.8, the expression for I_2 becomes:

$$\begin{aligned} I_2 &= \int_0^\eta q(Y_d)p(q(Y_d))dq(Y_d) \\ &= \int_{-\infty}^{\Phi^{-1}(\eta)} \Phi(z)f_q(\Phi(z))\phi(z)dz \end{aligned}$$

where $f_q(\Phi(z))$ is given by the derivative of $F_q = P(q(Y_d) < x)$ and $\phi(z)$ is the standard normal density:

$$\begin{aligned} P(q(Y_d) < x) &= P\left[\Phi\left(\frac{W_C - E_{Y_d}(X_{t+k})}{\mathrm{Var}_{Y_d}(X_{t+k})}\right) < x\right] \\ &= P(\Phi(z) < x) \\ &= P(z < \Phi^{-1}(x)) \\ &= P\left(\frac{z-\mu_z}{\sigma_z} < \frac{\Phi^{-1}(x)-\mu_z}{\sigma_z}\right) \\ &= \Phi\left(\frac{\Phi^{-1}(x)-\mu_z}{\sigma_z}\right) \end{aligned}$$

Therefore:

$$\begin{aligned} f_q = \frac{dF_q}{dx} &= \frac{d}{dx}\left[\Phi\left(\frac{\Phi^{-1}(x)-\mu_z}{\sigma_z}\right)\right] \\ &= \frac{1}{\sigma_z}\phi\left(\frac{\Phi^{-1}(x)-\mu_z}{\sigma_z}\right) \times \frac{1}{\phi(\Phi^{-1}(x))} \end{aligned}$$

Then:

$$\begin{aligned} I_2 &= \int_{-\infty}^{\Phi^{-1}(\eta)} \Phi(z) f_q(\Phi(z))\phi(z)dz \\ &= \int_{-\infty}^{\Phi^{-1}(\eta)} \Phi(z)\frac{1}{\sigma_z}\phi\left(\frac{\Phi^{-1}(\Phi(z))-\mu_z}{\sigma_z}\right) \times \frac{1}{\phi(\Phi^{-1}(\Phi(z)))}\phi(z)dz \\ &= \int_{-\infty}^{\Phi^{-1}(\eta)} \Phi(z)\frac{1}{\sigma_z}\phi\left(\frac{z-\mu_z}{\sigma_z}\right) \times \frac{1}{\phi(z)}\phi(z)dz \\ &= \int_{-\infty}^{\Phi^{-1}(\eta)} \Phi(z)\phi\left(\frac{z-\mu_z}{\sigma_z}\right)\frac{dz}{\sigma_z} \end{aligned}$$

These 2 integrals in the case of normality reduce to a closed form expression for I_1 and an integral I_2 which is easy to evaluate numerically. This means the calculation of the expected loss for a component is very quick, allowing us to quickly search the large design space.

8.5 Design selection

Total inspection cost incorporates expected loss from above, along with other costs associated with the process of carrying out inspections. For example, any inspection

will involve set-up costs. Inspection of some components will be more costly. Different designs might involve inspection of different numbers of components. Optimal designs should be selected with respect to total inspection cost, not only expected loss. To calculate the total loss for an inspection design, expected loss for each component is summed component-wise and added to the associated inspection cost. It is possible therefore to quantify the value of any design, d prior to carrying it out, and to search for good designs.

A method of searching efficiently for good designs from the space of designs is required. For example, even in the current simple case, with a binary inspection decision for each component, there are 2^C (C components) potential designs to choose from. Stepwise addition of components is one tractable search strategy; components are added sequentially to an empty starting design, such that at each step, the component added minimises the incremental total inspection cost. Alternatively a stepwise deletion, or any of a large number of possible search algorithms may be considered. Backward induction schemes for finding optimal designs might be useful but were not explored. In general a combination of stepwise addition and deletion works relatively well in practice.

8.6 Designing for variance learning

In section 8.2 we considered an expected loss criterion which preferentially chooses inspections which reduce model uncertainty. In doing so, the expected loss gets better predictions of the current mean system state and so reduces the risk of failure. The components within the model which are likely to have high uncertainty are those which we have not examined for a long time. Therefore the consequence of this type of criteria is that we will spread inspections out across many components and in time inspect the whole system.

It has been shown in section 5.3 that for variance updating, several observations of the same component are needed for inference. Therefore if we wanted to design inspection to best learn about system variance, then an optimal strategy would be to observe the same component at every inspection. The strategy for optimal variance

learning and the strategy for improving estimation of the system state are therefore contradictory. Can we design a utility criterion which attempts to do both?

8.6.1 Prior for $\mathcal{M}(W_X)$

Consider learning about the variances within the general effects corrosion model 3.5.1. In section 5.3 the adjusted mean variance, $E_D(\mathcal{M}(W_X))$ was computed to give an updated estimate for the general effects corrosion.

The adjusted expectation from equation 8.3.4 has prior expectation:

$$E_D(\mathcal{M}(W_X)) = E(\mathcal{M}(W_X)) = \Sigma_{W_X}$$

and from equation 8.3.5 variance:

$$\mathrm{Var}_D(\mathcal{M}(W_X)) = \mathrm{Var}(X_{t+k}) - \mathrm{Var}_{Y_d}(X_{t+k})$$

The following approach combines the design ideas in this chapter with the variance learning approach described in previous chapters to select good designs for simultaneous expectation and variance learning. Firstly we observe the system until time t and update beliefs about system mean and variances as described in sections 4.1, 5.6 and 7.5. Next choose a design d, and update beliefs about system to time $t+k$ given design d, using expectation and variance learning as described in section 8.2. Beliefs about the mean value of, $\mathcal{M}(W_X)$, given the design have mean, Σ_{W_X}, and variance, $\mathrm{Var}(X_{t+k}) - \mathrm{Var}_{Y_d}(X_{t+k})$, the resolved variance given design, d.

8.6.2 Gamma assumption

In a similar vein to the normality assumptions made in section 8.4 we specify a distributional form for $\mathcal{M}(W_X)$. We assume that $\mathcal{M}(W_X)$ takes a Γ distribution with mean, Σ_{W_X}, and variance, $\mathrm{Var}(X_{t+k}) - \mathrm{Var}_{Y_d}(X_{t+k})$. The choice of a Γ distribution for the variances enforces positivity of the variances. Given a mean and variance the Γ distribution can be parametrised thus:

$$\mathcal{M}(W_X) \sim \Gamma\left(\frac{\Sigma^2_{W_X}}{\mathrm{Var}(X_{t+k}) - \mathrm{Var}_{Y_d}(X_{t+k})}, \frac{\mathrm{Var}(X_{t+k}) - \mathrm{Var}_{Y_d}(X_{t+k})}{\Sigma_{W_X}}\right). \quad (8.6.11)$$

Let $f_d(\mathcal{M}(W_X))$ be the pdf of the Γ distribution specified in 8.6.11 given design d. To evaluate the expected loss of design d including the uncertainty around the variance we calculate the following integral:

$$E[L(O, \delta^*(Y_d))] = \int_0^\infty E[L(O, \delta^*(Y_d), \mathcal{M}(W_X))] f_d(\mathcal{M}(W)) d\mathcal{M}(W_X) \qquad (8.6.12)$$

where $E[L(O, \delta^*(Y_d), \mathcal{M}(W_X))]$ is the expected loss of the optimal design $\delta^*(Y_d)$ from equation, 8.2.3 for a given $\mathcal{M}(W)$.

Since we want to evaluate many design choices, we want to be able to compute this integral fast. This integral can be calculated numerically using a discretised version of the Γ distribution to weight the expected losses.

8.6.3 Numerical Approximation

Equation 8.6.12 can be approximated by:

$$\begin{aligned} E[L(O, \delta^*(Y_d))] &= \sum_{i=0}^{\infty} \int_i^{i+h} E[L(O, \delta^*(Y_d), \mathcal{M}(W_X))] f_d(\mathcal{M}(W_X)) d\mathcal{M}(W_X) \\ &\approx \sum_{i=0}^{\infty} hE[L(O, \delta^*(Y_d), \mathcal{M}(W))] \left(f_d(\frac{\mathcal{M}(W_{Xi}) + \mathcal{M}(W_{X(i+h)})}{2}) \right) \end{aligned}$$

As for the previous design criterion we can then search across the space of designs, d (as discussed in section 8.5), and choose the design which minimises total expected loss.

8.7 Example

8.7.1 Simulated example: Inspection and Maintenance

We will now consider an example of optimal inspection design using the expected loss criterion as described in section 8.4. Using synthetic data generated using corrosion model from section 3.5. We consider 4 components over 100 time points with each component inspected at 20 locations. Using the prior wall thickness and corrosion rate as follows:

$$X_0 = 100 \qquad\qquad \alpha_0 = -3$$

Every 5 time steps, we plan an inspection and decide component-wise whether or not to examine each component. Carrying out inspections has associated costs calculated as in equation 8.1.1 where the initial setup cost of an inspection is $L_{SUC} = 0.05$. The cost of inspecting a particular component is $L_C = 0.05$. If a component is inspected, and is below the critical wall thickness $W_C = 10$, then it is replaced.

System maintenance involves replacing a component, decision R, or leaving it alone, $\bar{R}$. Replacing a component incurs cost, $L_R = 1$, whereas component failure costs $L_F = 1000$ (so a component failure is considerably more expensive than replacement). If the corrosion rate were indeed 3 per time point then starting from an initial wall thickness of 100 we would expect a component to last ≈ 30 time points. So running over the 100 time points we would have expected to need to replace each component ≈ 3 times.

In this case there are only 16 designs (4^2) and at each time point we check all possible designs. For larger problems it would be impractical for check all designs. Figure 8.1 shows the resulting inspection and maintenance in this example. The first inspection doesn't take place until time 25 when component 4 is inspected. The "real" wall thickness at that time when observed is much higher than expected and so the predicted wall thickness is adjusted upwards and the estimated corrosion rate reduced. Due to the correlation in the system this inspection also causes the other non-inspected component wall thickness predictions to be adjusted upwards. The frequency of inspections increases as the components get close to the critical wall thickness. Maintenance of each of the 4 components occurs at times 50 and 55 without the need to inspect them all. Once maintenance has taken place then components are again inspected less frequency due to the decreased chance of component failure.

Over the whole time period 15 inspections take place on the 4 components with some components inspected more often than others. Due to the lower than expected corrosion rate the components in reality have a lifespan of ≈ 50 time points.

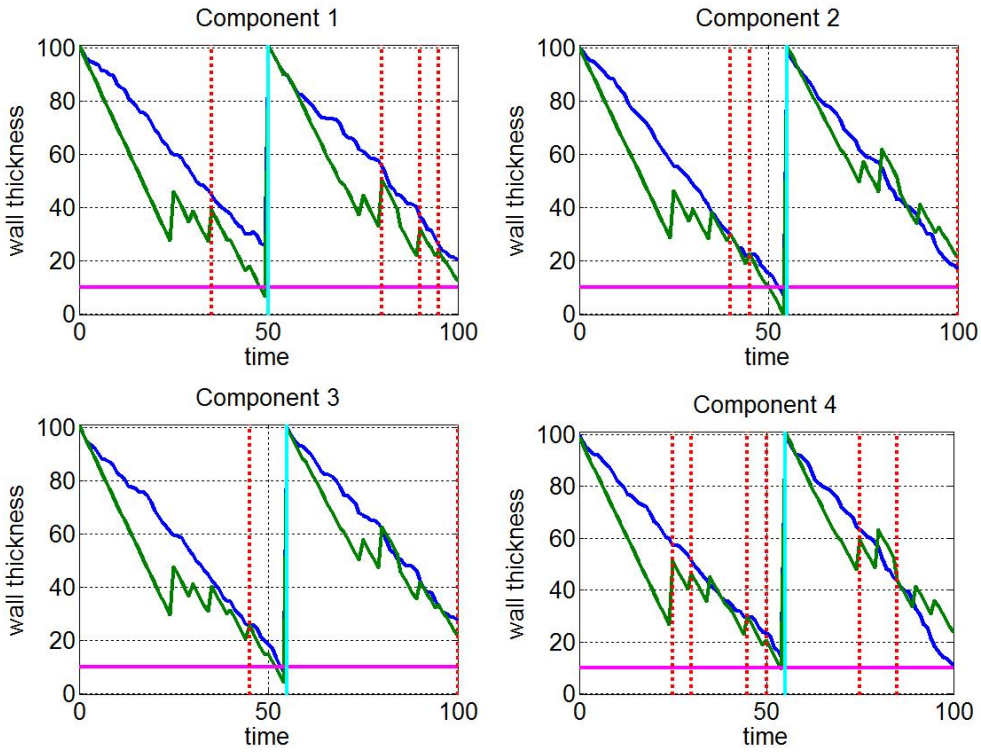

Figure 8.1: Inspection and maintenance of a 4 component system using synthetic data. The blue line shows the "real" component wall thickness at each time point. The green line shows the predicted wall thickness at any time. The horizontal magenta line shows the critical wall thickness below which the component must be replaced. The vertical dashed red line shows the times when inspections took place. The vertical cyan line shows the times maintenance took place.

8.7.2 Inspection Design with variance learning for offshore platform application

We continue the offshore platform example from previous chapters (previously discussed in sections 1.4.1, 3.5, 4.9, 5.7.1, 5.7.2, 6.4, 7.7 and 7.9.4).

We consider trying to design an inspection for the oil platform example in the case of analysis of the real inspection data described in section 1.4.1. A system of four corrosion circuits is modelled, consisting of a total of 64 pipe-work weld components, data is given in appendix A. We suppose we had just observed this data and updated the means and variances as discussed in sections 4.9 and 7.7.

We wish to decide how much of the system to inspect when trying to learn about means and variances as described in section 8.6.

In this example the cost of three different inspection designs are compared;

1. no inspection
2. inspection of half the system (every other component)
3. full inspection

In practice a large number of designs would be compared to try to find the optimal inspection scheme, this example is merely for illustration.

From section 7.8 the prior mean for $\mathcal{M}(W_X)$ is 0.12. Figure 8.2 shows the discretised Γ in the case of full inspection used to generate probabilities to weight expected loss estimates.

Critical System Components

Let each component have the same cost of replacement and failure:

$$L_R = 1 \qquad\qquad L_F = 100$$

so that the cost of component failure is 100 times the cost of replacement, representing vital system components. The cost of setting up an inspection is 0.01 per component inspected, so for inspection to be worthwhile the increased information about the system has to outweigh the increased cost.

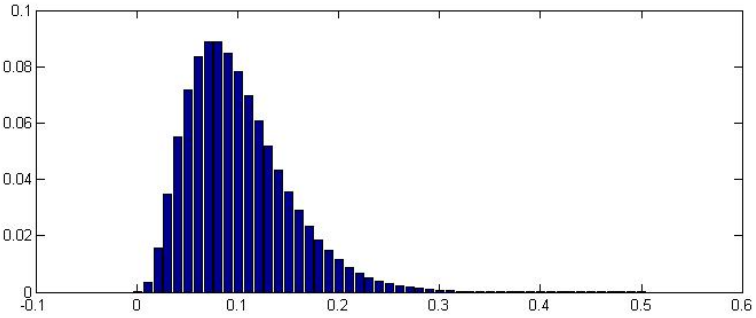

Figure 8.2: Discretised Γ distribution in the case of full inpsection for distribution of $\mathcal{M}(W_X)$

For this case:

	no insp.	half insp.	full insp.
Number of Components Inspected	0	32	64
Inspection Cost	0	0.32	0.64
Expected Loss (failures)	1.8774	1.598	**1.223**
Total Expected Loss	1.8774	1.918	**1.863**

We see that the expected loss from component failures goes down as we inspect more components since we have more information and less uncertainty. The expected loss from component failures must then be balanced against the cost of inspecting the system. In this case a full system inspection is the best, since in this cost of component failure relative to the cost of inspection is high.

Critical System Components: More Expensive Inspections

Lets consider the same case again but with the cost of inspection doubled to 0.02 per component.

	no insp.	half insp.	full insp.
Number of Components Inspected	0	32	64
Inspection Cost	0	0.64	1.28
Expected Loss (failures)	1.8774	1.598	**1.223**
Total Expected Loss	**1.8774**	2.238	2.503

Here the expected losses from failures are the same but the increased inspection costs has meant that no inspection is now the best choice.

Non Essential System Components

Lets consider another case where the cost of failure for each component is lower:

$$L_R = 1 \qquad\qquad L_F = 5$$

so that the cost of component failure is 5 times the cost of replacement. The cost of setting up an inspection is 0.01 per component inspected. In this case:

	no insp.	half insp.	full insp.
Number of Components Inspected	0	32	64
Inspection Cost	0	0.32	0.64
Expected Loss (failures)	1.042	0.935	**0.842**
Total Expected Loss	**1.042**	1.255	1.482

The expected loss from failures is lower than in previous cases since the risk of failure is lower. This means that with an inpsection cost of 0.01 per component the best design is no inspection.

Non Essential System Components: Cheaper Inspection

Consider the previous case this time with lower inspection costs of 0.003 per component. In this case:

	no insp.	half insp.	full insp.
Number of Components Inspected	0	32	64
Inspection Cost	0	0.096	0.192
Expected Loss (failures)	1.042	0.935	**0.842**
Total Expected Loss	1.042	**1.0310**	1.0340

Here with cheaper inspections we find that the best decision is to inspect half the system.

The expected loss criterion gives a method for comparing many different designs. The use of a utility based criterion also means than designs can be compared using

costs on a scale which an inspection planner can understand. The expected loss can incorporate both design tries to reduce uncertainty by preferentially choosing designs which improve mean and variance learning.

Chapter 9

Discussions and further study

This thesis has presented an approach to model large multivariate systems using Bayes linear updating to allow tractable inference. A model for system integrity was developed and it was shown how to use historical inspection data to update beliefs about system parameters, including both means and variances.

9.1 Modelling complex industrial systems

In chapter 1 the problem of complex industrial systems was introduced. These systems have high dimensionality and inspections of such systems are relatively infrequent and not carried out at regular intervals.

9.1.1 Model

A model was presented in chapter 3 which attempts to separate the characteristics of system components into to parts.

1. **General effects**: model average behaviour of system components and prescribes how components are related to each other through a potentially very large and difficult to specify dependency structure. The dependency allows the limited information for individual components to be shared around similarly behaving components.

2. **Local effects**: models the local surface deviations from the average component behaviour described by the general effects model. The degree of deviations depend on conditions local to each component.

Other work in this area usually treats components as independent entities using more complex modelling based around the physical process of corrosion. Often with the exception of some extremely critical system components, there is insufficient detailed inspection data to warrant this level of modelling detail.

The approach taken in the thesis was to produce a model which was flexible enough to describe system process but simple enough to allow fast updating of beliefs about the whole system. The advantage of this approach is in that information from short time series are combined and shared over the whole system. This gives better inference and the ability to learn about parts of the system which haven't ever been inspected.

9.1.2 Corrosion initiation

For the application within a corrosion setting, components will have measures in place to mitigate corrosion. Special coatings on components mean that in many cases corrosion would not begin from the moment a component is installed as would be predicted from the model. Corrosion engineers consider two distinct types of inspection;

1. Compliance sampling: to check whether corrosion was initiated, or whether there is measurable levels of corrosion present within a component.

2. Estimation sampling: once corrosion has begun this type of inspection tries to estimate corrosion rates and thus remaining life.

The modelling we have presented only really models the latter. One simple addition to the model would be to have a Poisson initiation time for each component before which the corrosion rate is zero. This minor alteration to the model would retain the simplicity needed to still generate fast realisations.

9.2 Updating model parameters

In chapter 4, an approach updating the current system state, was explored using Bayes Linear Analysis.

9.2.1 Full Bayesian

A full Bayesian approach to this type of problem would require full prior distributions specified in high dimensional space. Computation of posterior predictions for the system state need calculation of difficult integrals or alternatively a simulation based MCMC scheme which could take a long time to converge. In this type of modelling, simplifying assumptions in distributional forms are made to reduce computation complexity, even if the modeller thinks these assumptions are unrealistic. Add to this that we wish to evaluate potentially hundreds/thousands of inspection design choices and the full Bayesian approach becomes at best unappealing and very quickly intractable for real problems.

9.2.2 Bayes linear

Bayes linear analysis is different to a full Bayesian approach in that only partial priors need to be specified. Inference is carried out by computing adjusted expectations which are essentially systems of linear equations. This means that for complex systems, the prior specification and model inference remain tractable. Bayes Linear adjustment needs means, variances and covariances to be specified, to be able to adjust beliefs. For some quantities these too, can be difficult to specify, especially how dependencies evolve in time.

9.2.3 Simulation

To estimate the prior quantities needed for inference, a simulation approach is taken. The model used is simple enough to allow fast realisations to be generated for the whole system and its evolution in time. Once a set of realisations has been generated, simulation output can be used to calculate simple empirical estimates for any means, variances and covariances which are needed to do inference. If we do

enough simulations these empirical estimates encapsulate our prior beliefs and the model form. A MATLAB module based on this modelling, simulation and Bayes Linear updating approach has been developed for use in Shell's inspection software.

9.2.4 Prior specification

To begin generating model simulations some starting parameters still need to be specified. The most difficult of these starting parameters are the variances and covariance which control model evolution and dependencies. An approach to specifying correlation structures using relative proximity of components was used. The assumption being that the closer components are spatially, the more likely they are to exhibit similar characteristics. In reality, there is a more complex dependency structure; which would be very difficult to quantify. There is insufficient data to try to directly estimate this dependency, however we have learnt about its general characteristics.

9.3 Updating variance parameters

In chapters 5, 6 and 7 we explored methods for trying to learn about the variances within the model.

9.3.1 Bayes linear variance estimation

Second order exchangeability of observations in time and squared observations across components, was used to give a method of learning about the "mean" variability across several components at a time. To have any kind of estimate of variability a time series of at least 3 observations per component are required. Linear combinations of observations are used to remove system effects and give expressions containing only model residuals. Bayes linear adjusted expectation of the "mean" variance terms can then be used to update beliefs about all the variances within the model. These linear combinations mean that we are effectively learning about sums of several model variances at once. Confounding between model variances

make it difficult to disentangle their effects with the limited data available. As a consequence, ratios between model variances sometimes need to be fixed.

The assumptions of second order exchangeability greatly simplify modelling and could be extended to deal with even larger systems. In the corrosion context, we could think of exchangeable components, exchangeable corrosion circuits or even exchangeable oil platforms within the same model.

9.3.2 Mahalanobis variance learning

Some of the model variances are not possible to learn about directly using a combination of Bayes linear variance learning and exchangeability. Instead we adopt a method similar to profile likelihood of parameter estimation, using a candidate set of parameters and assess goodness of fit. A distance criterion is needed which contains information on the variance/covariance structure; a simple mean square error is insufficient. We use the Mahalanobis distance since this can describe the fit even in the situation of dependencies. The downside of a Mahalanobis distance is the lack of an estimate of variance to give an idea of model uncertainty. We can get an estimate in the case of a multivariate normal approximation. However, this is unlikely to hold in practice for the complex systems under consideration. Alternatively, synthetic data sets generated using updated model parameters can be used to explore model reliability.

9.3.3 Correlation updating

In section 7.9 we explored ways of extending variance estimation to learn about correlation. A cross validation style method was used, separating data into disjoint sets of separate components. In this way information can only be shared is through the correlation structure. Good model predictions therefore require a good understanding of the correlation structure. Given enough data this method would give an approach capable of learning about full covariance matrices. In practice using real data it was found found that there is insufficient information to reliably update correlation structures.

9.4 Design of inspections

9.4.1 Expected loss

Given methodology for modelling system integrity, and incorporating historical data, forecasts can be made. Simulations of the model can give predictions of current and future system status with associated uncertainties. Forecasts are inherently uncertain and can only give a guide. Inspections of the system are needed to give real measurements. The question arises; with limited time and money where are the optimal locations to inspect?

The locations we wish to observe are;

1. Components which have a high probability of failure. These are most likely to require maintenance.

2. Places about which we have large uncertainty. These could be components which we haven't observed for a long time, or which have weak dependency.

3. Places which improve our variance learning. The previous step would suggest we might want to examine components we haven't seen for a long time. Whereas for variance learning we might prefer to observe the same component many times to build up more information on its variability.

To try to address this question a utility criterion was used to balance the cost of information from inspection against the benefit from reduced uncertainty and chance of failures. Locations with high uncertainty or high risk of failure are to be prioritised. We can quickly evaluate the expected loss of a given inspection design. We can then search over potential designs and select the one which has the lowest expected loss. Using Bayes linear analysis with a model from which can get inference quickly allows us to evaluate a large number of designs.

The expected loss design approach could be extended in several ways. Inspections are not a one time thing and inspection planning should reflect that. Allowing components to be inspected at different times would allow this feature to be incorporated. It would seem likely that tuning upfront inspection costs would become more important to avoid inspecting small numbers of components at every time point.

There is an assumption of normality used making computation of the integrals in the expected loss fast. These assumptions are unappealing in the case of a minimum observation case since these are likely to be skewed. It is possible these assumptions could be relaxed without making the calculations intractable. Simulations are already needed to compute the adjusted expectations. These simulations contain information on the distribution of everything within the model. It seems likely that there is some way of using this information without the need to do further simulation.

Appendix A

Historical Inspection Data

Historical data for component minimum wall thickness, obtained during inspection campaigns using non-intrusive ultrasonic measurements for the period 1998 - 2005.

Corrosion Circuit	Component	Inspection Date	Observation
A	1	14/08/2001	6.35
A	1	30/08/2003	6.35
A	2	13/12/2000	6.35
A	2	14/08/2001	6.35
A	2	30/08/2003	6.35
A	3	01/06/1998	15
A	3	10/08/1999	15
A	3	13/12/2000	15
A	3	01/12/2001	18
A	3	06/12/2001	18
A	3	31/08/2003	15
A	4	10/08/1999	18.24
A	4	31/08/2003	17
A	5	13/12/2000	6.35
A	5	18/05/2003	6.35

continued on next page

continued from previous page

Corrosion Circuit	Component	Inspection Date	Observation
A	6	13/12/2000	6.35
A	6	14/08/2001	6.35
A	6	30/08/2003	6.35
A	7	13/12/2000	6.35
A	7	14/08/2001	6.35
A	7	30/08/2003	6.35
A	8	13/12/2000	4.78
A	8	14/08/2001	4.78
A	8	30/08/2003	6.35
A	9	10/08/1999	13.49
A	9	31/08/2003	9.5
A	9	07/07/2004	9.5
A	10	10/08/1999	13.49
A	10	12/08/2001	13.49
A	10	31/08/2003	9.5
A	10	07/07/2004	9
A	11	10/08/1999	7.1
A	11	18/05/2003	7.1
A	12	10/08/1999	7.1
A	12	18/05/2003	7.1
A	13	10/08/1999	5.56
A	13	18/05/2003	5.56
A	14	10/08/1999	5.56
A	14	12/08/2001	5.56
A	14	18/05/2003	5.56
A	15	10/08/1999	13.49
A	15	31/08/2003	9.5

continued on next page

continued from previous page

Corrosion Circuit	Component	Inspection Date	Observation
A	15	07/07/2004	9.5
A	16	10/08/1999	13.49
A	16	12/08/2001	13.49
A	16	31/08/2003	9
A	16	07/07/2004	9
A	17	10/08/1999	5.56
A	17	30/08/2003	3.7
A	17	07/07/2004	3.9
A	18	10/08/1999	5.56
A	18	12/08/2001	5.56
A	18	30/08/2003	6
A	19	10/08/1999	5.56
A	19	12/08/2001	5.56
A	19	30/08/2003	6
B	20	04/12/2000	7.14
B	20	01/12/2001	7.14
B	20	03/09/2003	8.1
B	21	04/12/2000	7.14
B	21	01/12/2001	7.14
B	21	03/09/2003	7.1
B	22	04/12/2000	6
B	22	01/12/2001	6
B	22	18/05/2003	5.35
B	23	04/12/2000	6
B	23	01/12/2001	6
B	23	18/05/2003	5.35
B	24	01/12/2001	4.78

continued on next page

continued from previous page

Corrosion Circuit	Component	Inspection Date	Observation
B	25	04/12/2000	4.78
B	25	15/08/2001	4.78
B	25	03/09/2003	5
B	26	03/09/2003	5
B	27	15/08/2001	8
B	27	16/08/2001	8
B	27	16/05/2003	7
B	27	30/08/2003	6
B	28	15/08/2001	12
B	28	16/08/2001	12
B	28	16/05/2003	11
B	28	30/08/2003	9.5
B	29	18/05/2003	1.91
B	30	18/05/2003	3.41
B	31	02/09/2003	20
B	32	02/09/2003	20
B	33	02/09/2003	20
B	34	02/09/2003	21
C	35	10/12/1998	17
C	35	12/08/2001	16
C	35	18/08/2001	16
C	35	14/05/2003	16
C	36	10/12/1998	6.35
C	36	09/12/2000	6
C	36	18/05/2003	5.35
C	37	10/12/1998	6.35
C	37	09/12/2000	6

continued on next page

continued from previous page

Corrosion Circuit	Component	Inspection Date	Observation
C	37	18/05/2003	5.35
C	37	04/09/2003	6.35
C	38	10/12/1998	13
C	38	07/05/2000	15
C	38	09/12/2000	12
C	38	12/08/2001	13.5
C	38	18/08/2001	13.5
C	38	14/05/2003	13
C	38	03/09/2003	8
C	38	08/04/2004	8
C	38	20/05/2004	8
C	39	10/12/1998	6.35
C	39	18/05/2003	3.85
C	39	04/09/2003	5.8
C	40	10/12/1998	6.35
C	40	09/12/2000	6.35
C	40	04/09/2003	6.35
C	41	10/12/1998	6.35
C	41	09/12/2000	6.35
C	41	18/05/2003	4.85
C	41	04/09/2003	4.5
C	41	08/04/2004	5.5
C	42	10/12/1998	6.35
C	42	09/12/2000	6.35
C	42	18/05/2003	6.35
C	43	10/12/1998	12
C	43	09/12/2000	9.5

continued on next page

continued from previous page

Corrosion Circuit	Component	Inspection Date	Observation
C	43	12/08/2001	11
C	43	18/08/2001	11
C	43	14/05/2003	9.5
C	43	03/09/2003	8
C	43	08/04/2004	8
C	44	10/12/1998	13.49
C	44	03/09/2003	10.5
C	44	08/04/2004	10.5
C	45	10/12/1998	5
C	45	09/12/2000	5
C	45	12/08/2001	7.1
C	45	18/05/2003	7.1
C	46	10/12/1998	4
C	46	09/12/2000	5.56
C	46	18/05/2003	3.56
C	47	10/12/1998	5
C	47	09/12/2000	5.56
C	47	18/05/2003	5.56
C	48	10/12/1998	13.49
C	48	03/09/2003	10
C	48	08/04/2004	10.5
C	49	10/12/1998	13.49
C	49	03/09/2003	10.5
C	49	08/04/2004	10
C	50	10/12/1998	5.56
C	50	18/05/2003	3.56
C	50	04/09/2003	3.9

continued on next page

continued from previous page

Corrosion Circuit	Component	Inspection Date	Observation
C	51	10/12/1998	5.56
C	51	18/05/2003	5.56
D	52	12/06/1998	6.35
D	52	12/12/2000	5.35
D	52	18/05/2003	5.35
D	53	12/06/1998	6.35
D	53	12/12/2000	5.35
D	53	18/05/2003	5.35
D	53	31/08/2003	6.35
D	54	12/06/1998	4.78
D	54	12/12/2000	4.78
D	54	18/05/2003	4.78
D	55	12/06/1998	8.74
D	55	12/12/2000	8.74
D	55	18/05/2003	8.74
D	56	17/05/2003	12
D	57	12/06/1998	7.1
D	57	12/12/2000	7.1
D	57	31/08/2003	6.35
D	58	12/12/2000	13
D	58	17/05/2003	13
D	59	18/05/2003	5.56
D	60	12/12/2000	5.56
D	60	18/05/2003	5.56
D	61	17/05/2003	14
D	62	12/12/2000	7.14
D	63	18/05/2003	3.56

continued on next page

continued from previous page			
Corrosion Circuit	Component	Inspection Date	Observation
D	63	31/08/2003	2.6
D	63	08/04/2004	2.8
D	64	12/12/2000	5
D	64	18/05/2003	5.56

Table A.1: Corrosion Data, (This table is split across pages)

Appendix B

Bayes Linear Variance learning for corrosion model

B.1 $\mathrm{Cov}[\mathcal{M}(W_X), D]$

$$\mathrm{Cov}[\mathcal{M}(W_X), D] = (\mathrm{Cov}[\mathcal{M}(W_X), D_1], \mathrm{Cov}[\mathcal{M}(W_X), D_2], \ldots, \mathrm{Cov}[\mathcal{M}(W_X), D_n])$$

and from equation 6.2.4,

$$D_{ct} = (k_i Y_c^{(l_i)} - l_i Y_c^{(k_i)})^2$$

Using equations in section 5.5.2

$$\begin{aligned}(k_i Y_c^{(l_i)} - l_i Y_c^{(k_i)})^2 = &\begin{pmatrix} (k_i - l_i) & 0 \end{pmatrix} \xi_{c(t_i, t_i - k_i)} + \begin{pmatrix} k_i & k_i(k_i - l_i) \end{pmatrix} \xi_{c(t_i - k_i, t_i - l_i)} \\ &+ \left(k_i M_c^{(l_i)} - l_i M_c^{(k_i)}\right)^2\end{aligned}$$

which means

$$\mathrm{Cov}\left[\mathcal{M}(W_X), (k_i Y_c^{(l_i)} - l_i Y_c^{(k_i)})^2\right] =$$

$$\begin{aligned}&= \mathrm{Cov}\left[\mathcal{M}(W_X), \left(\begin{pmatrix} (k_i - l_i) & 0 \end{pmatrix} \xi_{c(t_i, t_i - k_i)}\right)^2\right] \\ &+ \mathrm{Cov}\left[\mathcal{M}(W_X), \left(\begin{pmatrix} k_i & k_i(k_i - l_i) \end{pmatrix} \xi_{c(t_i - k_i, t_i - l_i)}\right)^2\right]\end{aligned}$$

we will consider

A.

$$\mathrm{Cov}\left[\mathcal{M}(W_X), \left(\begin{pmatrix} (k_i - l_i) & 0 \end{pmatrix} \xi_{c(t_i, t_i - k_i)}\right)^2\right]$$

B.

$$\mathrm{Cov}\left[\mathcal{M}(W_X), \left(\begin{pmatrix} k_i & k_i(k_i-l_i) \end{pmatrix} \xi_{c(t_i-k_i,t_i-l_i)}\right)^2\right]$$

B.1.1 A

$$\begin{aligned}
& \mathrm{Cov}\left[\mathcal{M}(W_X), \left(\begin{pmatrix} (k_i-l_i) & 0 \end{pmatrix} \xi_{c(t_i,t_i-k_i)}\right)^2\right] \\
= \; & \mathrm{Cov}\left[\mathcal{M}(W_X), \left(\sum_{j=0}^{k_i-1} \begin{pmatrix} (k_i-l_i) & 0 \end{pmatrix} \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_X+\epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{ct_i-j}\right)^2\right] \\
= \; & \mathrm{Cov}\left[\mathcal{M}(W_X), \left(\sum_{j=0}^{k_i-1} \begin{pmatrix} (k_i-l_i) & j(k_i-l_i) \end{pmatrix} \begin{pmatrix} \epsilon_X+\epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{ct_i-j}\right)^2\right] \\
= \; & \mathrm{Cov}\left[\mathcal{M}(W_X), \sum_{j=0}^{k_i-1}\sum_{j'=0}^{k_i-1} \begin{array}{l} (k_i-l_i)^2\epsilon_{Xct_i-j}\epsilon_{Xct_i-j'} \\ +(j'+1)(k_i-l_i)^2\epsilon_{Xct_i-j}\epsilon_{\alpha ct_i-j'}+ \\ +(j'+1)(k_i-l_i)^2\epsilon_{Xct_i-j}\epsilon_{\alpha ct_i-j'}+ \\ +(j'+1)(j+1)(k_i-l_i)^2\epsilon_{\alpha ct_i-j}\epsilon_{\alpha ct_i-j'} \end{array}\right] \\
= \; & \mathrm{Cov}\left[\mathcal{M}(W_X), \sum_{j=0}^{k_i-1}(k_i-l_i)^2\epsilon^2_{Xct_i-j}\right] \\
= \; & (k_i-l_i)^2\sum_{j=0}^{k_i-1}\mathrm{Cov}\left[\mathcal{M}(W_X), \epsilon^2_{Xct_i-j}\right] \\
= \; & (k_i-l_i)^2\sum_{j=0}^{k_i-1}\mathrm{Cov}\left[\mathcal{M}(W_X), \mathcal{M}(W_X)+\mathcal{R}_c(W_X)+\mathcal{R}_t(V_{Xc})\right] \\
= \; & (k_i-l_i)^2\sum_{j=0}^{k_i-1}\mathrm{Var}\left[\mathcal{M}(W_X)\right] \\
= \; & k_i(k_i-l_i)^2\Xi_{W_X}
\end{aligned}$$

B.1.2 B

$$
\begin{aligned}
& \mathrm{Cov}\left[\mathcal{M}(W_X), \left(\begin{pmatrix} k_i & k_i(k_i-l_i) \end{pmatrix} \xi_{c(t_i-k_i,t_i-l_i)}\right)^2\right] \\
= \; & \mathrm{Cov}\left[\mathcal{M}(W_X),\right. \\
& \left.\left(\sum_{j=0}^{l_i-k_i-1} \begin{pmatrix} k_i & k_i(k_i-l_i) \end{pmatrix} \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{c(t_i-k_i-j)}\right)^2\right] \\
= \; & \mathrm{Cov}\left[\mathcal{M}(W_X),\right. \\
& \left.\left(\sum_{j=0}^{l_i-k_i-1} \begin{pmatrix} k_i & jk_i + k_i(k_i-l_i) \end{pmatrix} \begin{pmatrix} \epsilon_X + \epsilon_\alpha \\ \epsilon_\alpha \end{pmatrix}_{c(t_i-k_i-j)}\right)^2\right] \\
= \; & \mathrm{Cov}\left[\mathcal{M}(W_X), \left(\sum_{j=0}^{l_i-k_i-1} k_i\epsilon_{Xct_i-k_i-j} + k_i(j+1+k_i-l_i)\epsilon_{\alpha ct_i-k_i-j}\right)^2\right] \\
= \; & \mathrm{Cov}\left[\mathcal{M}(W_X),\right. \\
& \left.\sum_{j=0}^{l_i-k_i-1}\sum_{j'=0}^{l_i-k_i-1} \begin{array}{l} \left(k_i\epsilon_{Xct_i-k_i-j} + k_i(j+1+k_i-l_i)\epsilon_{\alpha c(t_i-k_i-j)}\right) \\ \times \left(k_i\epsilon_{Xct_i-k_i-j'} + k_i(j'+1+k_i-l_i)\epsilon_{\alpha c(t_i-k_i-j')}\right) \end{array}\right] \\
= \; & \mathrm{Cov}\left[\mathcal{M}(W_X), k_i^2 \sum_{j=0}^{l_i-k_i-1} \epsilon^2_{Xc(t_i-k_i-j)}\right] \\
= \; & (l_i-k_i)k_i^2\Xi_{W_X}
\end{aligned}
$$

Then

$$
\begin{aligned}
\mathrm{Cov}\left[\mathcal{M}(W_X), (k_iY_c^{(l_i)} - l_iY_c^{(k_i)})^2\right] &= \mathbf{A} + \mathbf{B}. \\
&= k_i(k_i-l_i)^2\Xi_{W_X} + (l_i-k_i)k_i^2\Xi_{W_X} \\
&= k_il_i(l_i-k_i)\Xi_{W_X}
\end{aligned}
$$

Appendix C

Table of Notation

Symbol	Description	Format	Defined
A	r.q.	r.q.	5.1
c	component index	scalar index	1.4
C	numbers of components in the model	scalar	3.1.2
Cir	corrosion circuit	scalar	4.5.3
$\text{Dis}_Y(\Theta_T)$	adjustment discrepancy of quantity Y	scalar	4.7.5
$\text{Dis}_Y(\Theta_T)$	adjustment discrepancy ratio	scalar	4.7.5
$\text{Dis}(Y)$	discrepancy of quantity Y	scalar	4.7.2
$\text{Dr}(Y)$	discrepancy ratio Y	scalar	4.7.2
d	inspection design	vector	8.2
D	data vector	vector	5.1.3
D_{ct}	squared linear combination of observations used in variance learning	scalar	5.3.4
D_c	vector in time of squared linear combination of observations used in variance learning	vector	5.3.4
$\bar{D_T}$	average of differences of observations D	scalar	5.2
$\bar{D}_n^{(2)}$	average of squared observations D	scalar	5.1.3
f	non linear observation function over locations	functional	3.3.2

continued on next page

continued from previous page

Symbol	Description	Format	Defined
F	outcome component failure	outcome	8.1
$\bar{F}$	outcome component survival	outcome	8.1
F_c	c'th row of F the dynamic regression matrix	vector	3.3.2
F	dynamic regression matrix	matrix	3.1.1
F_c^*	c'th row of F^* the dynamic regression matrix	vector	3.3.2
F^*	dynamic regression matrix (outside observation function)	matrix	3.3.2
g	non linear function describing the local surface	functional	3.2.1
G	system state evolution matrix	matrix	3.1.1
k_i	time between observations t_i and $t_(i-1)$	scalar	5.5.2
l_i	time between observations t_i and $t_(i-2)$	scalar	5.5.2
I_1	Integral used in calculation of expected loss	integral	8.2
I_2	Integral used in calculation of expected loss	integral	8.2
l	location index on the surface of a component	scalar index	3
L_c	number of locations observed on a component c	scalar	3
L_F	Cost associated with component failure	scalar	8.1
L_IC	Cost of whole inspection campaign	scalar	8.1
L_R	Cost associated with component replacement	scalar	8.1
L_{SUC}	Cost of setting up an inspection campaign	scalar	8.1
M_{ct}	minima of local effects model	scalar	6.2.1
$M_c^{(i)}$	ith time step differences of minima M	scalar	6.2.1
$\mathcal{M}(X)$	"population mean" of X in the representation theorem	r.q.	2.3.2
$\mathcal{M}_\Theta$	"population mean" for Θ	r.q.	4.5.1
n	parameter index	scalar index	3.1.1

continued on next page

continued from previous page

Symbol	Description	Format	Defined
n_d	number of inspected components	scalar	8.1
N	number of parameters in the model	scalar	3.1.1
o	particular outcome	outcome	8.1
O	Outcome space	outcome	8.1
p	parameter index subset of full parameter space	scalar index	3.1.3
P	size of subset of parameter space $P \leq N$	scalar	3.1.3
$q(Y_d)$	probability of failure	scalar	8.3
r_{lt}	vector of local effects over components at location l, at time t,	vector	3.3.1
r_{lct}	local effects at location l, on component c, at time t	scalar	3.2.1
R	decision to replace component	decision	8.1
$\bar{R}$	decision not to replace component	decision	8.1
$\mathcal{R}_i(X)$	"population residual" of X vector in representation theorem	r.q.	2.3.2
$\mathcal{R}_{\Theta ct}$	"population residual" for Θ	r.q.	4.5.1
S	adjacency/distance matrix	matrix	4.5.2
$S_{cc'}$	distance between components c and c'	scalar	4.5.2
t	time index	scalar index	1.4.1
t_i	time of ith observation	scalar	5.5.2
T	number of time points in model	scalar	4.1
$V_{\alpha ct}$	the squared corrosion rate evolution residuals $\epsilon^2_{\alpha ct}$ for component c, at time t	r.q.	3.5.1
V_{nt}	the squared evolution residuals $\epsilon^2_{\Theta nt}$ for parameter n, at time t	r.q.	3.1.3
V_{Xct}	the squared wall thickness evolution residuals ϵ^2_{Xct} for component c, at time t	r.q.	3.5.1

continued on next page

continued from previous page

Symbol	Description	Format	Defined
W_C	Critical wall thickness	scalar	8.3
$W_{\alpha c}$	$\mathcal{M}(V_{\alpha c})$ for component c, at time t	r.q.	3.5.1
W_{Xc}	$\mathcal{M}(V_{Xc})$ for component c, at time t	r.q.	3.5.1
W_p	$\mathcal{M}(V_p)$ for parameter p, at time t	r.q.	3.1.3
X_{ct}	wall thickness of component c at time t	scalar	3.5.1
X_t	wall thickness vector over components at time t	vector	3.5.1
X_{ct}	wall thickness of component c at time t	scalar	5.3
Y_{ct}	observation on component c at time t	scalar	3.3.2
$Y_{ct}^{(i)}$	ith time step differences of observations Y	scalar	5.3.3
Y	vector of all observations for full inspections	vector	3.3.2
Z_t	vector of true system state over components at time t	vector	3.1.1
Z_{lt}	vector over components of true system state at time t and location l	vector	3.3.1
Z_{lct}	true system state at location l, component c, and time t	scalar	3.3.2
α_{ct}	corrosion rate of component c at time t	scalar	3.5.1
α_t	corrosion rate vector over components at time t	vector	3.5.1
γ	covariance parameter	scalar	7.3
γ_Θ	$\mathrm{Cov}(\epsilon_{\Theta ct}, \epsilon_{\Theta c't'})$ in exchangeability representation	scalar	4.5.1
$\delta_{\mathrm{CirCir}'}$	1 if same circuit; 0 if different circuits	scalar	4.5.3
δ^*	optimal decision	decision	8.2
Δ	space of decisions	decision	8.2
$\epsilon_{\alpha ct}$	corrosion rate evolution residual over parameters on component cat time t	scalar	3.5.1

continued on next page

continued from previous page

Symbol	Description	Format	Defined
ϵ_{rlct}	local effects residual at location l, on component c, at time t	scalar	3.2.1
$\epsilon_{\Theta t}$	system state evolution residual vector over parameters at time t	vector	3.1.1
$\epsilon_{\Theta nt}$	system state evolution residual for parameter n, at time t	scalar	3.1.3
ϵ_{Xct}	wall thickness evolution residual over parameters on component c at time t	scalar	3.5.1
ϵ_{Ylct}	observation error residual at location l, component, c at time t	scalar	3.3.2
η	ratio of cost of replace to cost of failure	scalar	8.2
Θ_t	system state vector over parameters at time t	vector	3.1.1
Θ_{nt}	system state for parameter n at time t	scalar	3.1.1
Θ_{ct}	system state for component c at time t	scalar	3.1.1
μ_A	expectation of r.q. A	scalar	5.1
ν	exponential decay term in adjacency matrix	scalar	4.5.2
$\xi_{c(t,t-k)}$	generalised residual for component c within generalised DLM	vector	5.5.1
$\Xi_{V_{Xc}}$	4th order quantity $\mathrm{Var}(\mathcal{M}(V_{Xc}))$	scalar	5.3.2
$\Xi_{V_{\alpha c}}$	4th order quantity $\mathrm{Var}(\mathcal{M}(V_{\alpha c}))$	scalar	5.3.2
$\Xi_{V_{Yc}}$	4th order quantity $\mathrm{Var}(\mathcal{M}(V_{Yc}))$	scalar	5.3.2
Ξ_{W_α}	4th order quantity $\mathrm{Cov}(W_{\alpha c}, W_{\alpha c'})$ where $c \neq c'$	scalar	3.5.1
Ξ_{W_X}	4th order quantity $\mathrm{Cov}(W_{Xc}, W_{Xc'})$ where $c \neq c'$	scalar	3.5.1
Ξ_W	4th order quantity $\mathrm{Cov}(W_p, W_{p'})$ where $p \neq p'$	scalar	3.1.3

continued on next page

continued from previous page

Symbol	Description	Format	Defined
ρ	correlation parameter	scalar	7.3
ρ_{Cir}	circuit correlation	scalar	4.5.3
ρ_0	underlying universal correlation	scalar	4.5.3
$\Sigma_{\alpha c}$	variance for α of component c	scalar	3.5.1
$\Sigma_{\alpha nn'}$	$\text{Cov}[\epsilon_{\alpha ct}, \epsilon_{\alpha ct}]$ with $c \neq c'$	scalar	3.5.1
Σ_A	variance of r.q. A	scalar	5.1
Σ_r	local effects variance parameter $\text{Var}(\epsilon_{rlct})$	matrix	3.2.2
Σ_Θ	variance matrix for Θ parameterised by σ^2_Θ and γ_Θ	matrix	4.5.1
$\sigma_{\Theta n}$	variance for $\epsilon_{\Theta nt}$ of system state parameter n	scalar	4.5.2
$\Sigma_{\Theta nn'}$	covariance for Θ between system state parameter n and n'	scalar	4.5.2
σ^2_Θ	$\text{Var}(\epsilon_{\Theta ct})$ in exchangeability representation	scalar	4.5.1
Σ_{W_α}	expected value, $E(W_{\alpha c})$	scalar	3.5.1
Σ_{W_X}	expected value, $E(W_{Xc})$	scalar	3.5.1
Σ_W	expected value, $E(W_p)$	scalar	3.1.3
$\Sigma_{\Theta n}$	$\text{Var}(\epsilon_{\Theta nt}) = \Sigma_{\Theta n}$	scalar	3.1.3
$\Sigma_{\Theta nn'}$	$\text{Cov}[\epsilon_{\Theta nt}, \epsilon_{\Theta n't}]$ with $n \neq n'$	scalar	3.1.3
Σ_Θ	$N \times N$ system state evolution variance matrix	matrix	3.1.1
Σ_X	variance of wall thickness in DLM	scalar	5.2
Σ_{Xc}	$\text{Var}(\epsilon_{Xct}) = \Sigma_{Xc}$ wall thickness variance	scalar	3.5.1
$\Sigma_{Xcc'}$	$\text{Cov}[\epsilon_{Xct}, \epsilon_{Xc't}]$ with $c \neq c'$ wall thickness covariance	scalar	3.5.1
Σ_Y	observation variance matrix	matrix	3.3.2
Σ_{Yc}	$\text{Var}[\epsilon_{Xct}]$ observation error variance	scalar	5.3.1
$\Sigma_{Ycc'}$	$\text{Cov}[\epsilon_{Yct}, \epsilon_{Yc't}]$ with $c \neq c'$	scalar	5.3.1
Φ_V	4th order quantity of $\text{Var}(V_i)$	scalar	5.1
Φ_{V_n}	4th order quantity of squared evolution residuals $\text{Var}(V_{nt})$	scalar	3.1.3

continued on next page

continued from previous page

Symbol	Description	Format	Defined
$\Phi_{V_{Xc}}$	4th order quantity $\mathrm{Var}(V_{Xc})$	scalar	5.3.2
$\Phi_{V_{\alpha c}}$	4th order quantity $\mathrm{Var}(V_{\alpha c})$	scalar	5.3.2
$\Phi_{V_{Yc}}$	4th order quantity $\mathrm{Var}(V_{Yc})$	scalar	5.3.2
Φ_{W_α}	4th order quantity of $\mathrm{Var}(W_{\alpha c})$	scalar	3.5.1
Φ_{W_X}	4th order quantity of $\mathrm{Var}(W_{Xc})$	scalar	3.5.1
Φ_W	4th order quantity of $\mathrm{Var}(W_p)$	scalar	3.1.3
$\dagger$	inverse if invertible else generalised inverse	function	2.2.1

Table C.1: Table of Notation, (This table is split across pages)

Bibliography

ASNT. Introduction to nondestructive testing. *American Society for Nondestructive Testing*, 2009. URL `http://www.asnt.org/ndt/primer2.htm`.

ASTM-Standard-G16. *Standard guide for applying statistics to analysis of corrosion data*, 2004. URL `http://www.npl.co.uk/upload/pdf/beginners_guide_to_corrosion.pdf`.

F. Caleyo, J. C. Velazquez, A. Valor, and J. M. Hallen. Probability distribution of pitting corrosion depth and rate in underground pipelines: A Monte Carlo study. *CORROSION SCIENCE*, 51(9):1925–1934, 2009.

S. Coles. *An introduction to statistical modeling of extreme values*. Springer, London, 2001.

P.S. Craig, M. Goldstein, J.C. Rougier, and A.H. Seheult. Pressure matching for hydrocarbon reserviors: A case study in the use of bayes linear strategies for large computer experiments. In C. Gatsonis, J.S. Hodges, R.E. Kass, R. McCulloch, P. Rossi, and N.D. Singpurwall, editors, *Case Studies in Bayesian Statistics: Volume III*, pages 36–93. Springer-Verlag, New York, 1997.

P.S. Craig, M. Goldstein, J.C. Rougier, and A.H. Seheult. Bayesian forecasting for complex systems using computer simulators. *American Statistical Association*, 96, 2001.

B. de Finetti. *Theory of Probability*. Wiley series in probability and mathematical statistics. Wiley, London, 1974.

P. Diggle and A. Verbyla. Nonparametric estimation of covariance structure in longitudinal data. *Biometrics*, 54:401–415, 1998.

A.C. Estes and D.M. Frangopol. Minimum expected cost-oriented optimal maintenance planning for deteriorating structures. *Reliability Engineering and System Safety*, 73:281–289, 2001.

M. H. Faber and J.D. Sorensen. Indicators for assessment and inspection planning. *Proceedings ESRA Workshop on Reliability and Risk Based Inspection Planning*, 2000.

M. Farrow. Practical building of subjective covariance structures for large complicated systems. *Statistician*, 52:553–573, 2003.

M. Farrow and M. Goldstein. Bayes linear methods for grouped mulitvariate repeated measurement studies with application to crossover trials. *Biometrika*, 80: 39–59, 1993.

J. Gasemyr and B. Natvig. Bayesian inference based on partial monitoring of components with applications to preventive system maintenance. *Naval Research Logisitics*, 48:551–577, 2001.

G.Hardman. *Bayesian Inspection Planning for large industrial Systems*. PhD thesis, Durham University, 2007.

M. Glegola. *Extreme Value Analysis of Corrosion Data*, 2007.

M. Goldstein. Exchangeable belief structures. *Journal of the American Statistical Association*, 81:971–976, 1986.

M. Goldstein. Bayes linear analysis. *Encyclopaedia of Statistical Sciences*, 3, 1998.

M. Goldstein and J.C. Rougier. Bayes linear calibrated prediction for complex systems. *American Statistical Association*, 101, 2006.

M. Goldstein and D. Wooff. *Bayes Linear Statistics: Theory and Methods*. Wiley Series in Probability and Statistics. Wiley, New York USA, 2007.

M Goldstein and DA Wooff. Adjusting exchangeable beliefs. *BIOMETRIKA*, 85: 39–54, 1998.

J. Harrison and M. West. *Bayesian Forecasting and Dynamic Models*. Springer Series in Statistics. Springer-Verlag, New York USA, 1st edition, 1989.

D. Hawn. Extreme value prediction of maximum pits on pipelines. *MATERIALS PERFORMANCE*, 16(3):29–32, 1977.

UK Health and Safety Executive. *Guidelines for use of statistics for analysis of sample inspection of corrosion*, 2002.

J. P. Imhof. Computing the distribution of quadratic forms in normal variables. *Biometrika*, 46:419–426, 1961.

MJ Kallen and JM van Noortwijk. Optimal maintenance decisions under imperfect inspection. *RELIABILITY ENGINEERING & SYSTEM SAFETY*, 90(2-3): 177–185, 2005. 14th European Safety and Reliability Conference (ESREL 2003), MAASTRICHT, NETHERLANDS, JUN 15-18, 2003.

S. Kotz and S. Nadarajah. *Extreme Value Distributions*. Imperial College Press, London, 2000.

J. Krautkrämer and H. J. Krautkrämer. *Ultrasonic Testing of Materials*. Springer Velag, 5th edition, 1990.

Sebastian P. Kuniewski, Johannes A.M. van der Weide, and Jan M. van Noortwijk. Sampling inspection for the evaluation of time-dependent reliability of deteriorating systems under imperfect defect detection. *Reliability Engineering and System Safety*, 94(9):1480 – 1490, 2009. ISSN 0951-8320. doi: 10.1016/j.ress.2008.11.013. URL `http://www.sciencedirect.com/science/article/pii/S0951832008002718`. ¡ce:title¿ESREL 2007, the 18th European Safety and Reliability Conference¡/ce:title¿.

J. Little, M. Goldstein, and P. Jonathan. Efficient bayesian sampling inspection for industrial processes based on transform spatio-temporal data. *Statistical Modelling*, 4:299–313, 2004a.

J. Little, M. Goldstein, P. Jonathan, and K. den Heijer. Spatio-temporal modelling of corrosion in an industrial furnace. *Appl. Stochastic Models Bus. Ind.*, 20:219–238, 2004b.

H. Liu, Y. Tang, and H. Zhang. A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables. *Computational Statistics and Data Analysis*, 53:853–856, 2009.

P.C. Mahalanobis. On the generalized distance in statistics. *Proceedings of the National Institute of Science of India*, 12:49–55, 1936.

B. Nimmo and Gareth Hinds. Begginers guide to corrosion. *National Physical Laboratory*, 2003. URL `http://www.npl.co.uk/upload/pdf/beginners_guide_to_corrosion.pdf`.

A. O'Hagan, C. Buck, A. Daneshkhah, J. Eiser, P. Garthwaite, D. Jenkinson, J. Oakley, and T. Rakow. *Uncertain Judgements: Eliciting Expert's Probabilities*. Statistics in Practice. Wiley, New York USA, 2007.

F. Pukelsheim. The 3-Sigma-Rule. *AMERICAN STATISTICIAN*, 48(2):88–91, 1994.

S. Qin and W. Cui. Effect of corrosion models on the time-dependent reliability of steel plated elements. *Marine Structures*, 16:15–34, 2003.

H. Raiffa and R. Schlaifer. *Applied Statistical Decision Theory*. Division of Research, Graduate School of Business Administration, Harvard University, 1961.

D. Randell, M. Goldstein, Hardman G, and P. Jonathan. Bayesian linear inspection planning for large-scale physical systems. *Journal of Risk and Reliability*, 224: 333–345, 2010.

M. Y. Rawwash. Covariance matrix estimation using repeated measurements when data are incomplete. *Applied Mathematics and Computation*, 167:299–315, 2005.

C.P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer-Verlag, New York USA, 2nd edition, 2004.

P. Scarf and P. Laycock. Estimation of extremes in corrosion engineering. *JOURNAL OF APPLIED STATISTICS*, 23(6):621–643, 1996.

G. Shaddick and J. Wakefield. Modelling daily multivariate pollutant data at multiple sites. *Appl. Statist.*, 51:351–372, 2002.

S. Shaw. *Simplfying Bayesian Experimental Design for multivariate partially exchangeable systems*. PhD thesis, University of Durham, 2000.

D. Straub. *Generic approaches to Risk Based Inspection Planning for Steel structures*. PhD thesis, Swiss Federal Institute of Technology Zurich, 2004.

D. Talbot and J. Talbot. *Corrosion Science and Technology*. CRC, 1998.

H. Uhlig. *Corrosion and Corrosion Control*. Wiley, 2nd edition, 1971.

D. F. Vysochanskij and Y. I. Petunin. Justification of the 3σ rule for unimodal distributions. *Theory of Probability and Mathematical Statistics*, 21:2536, 1980.

P. Whittle. *Probability via Expectation*. Springer-Verlag, New York, 3rd edition, 1992.

D. Wilkinson and M. Goldstein. Bayes linear covariance matrix adjustment for multivariate dynamic linear models, 1997. URL `citeseer.ist.psu.edu/172077.html`.

D. J. Wilkinson. Bayes linear variance adjustment for locally linear dlms. *Journal of Forecasting*, 16:329–342, 1997.

W. B. Wu and M Pouramadi. Nonparametric estimation of large covariance matrices of longitudinal data. *Biometrika*, 90:831–844, 2003.

R. Zhang and S. Mahadevan. Ndi-based corrosion reliability reassment. *8th ASCE Speciality Conference on Probabilistic Mechanics and Structural Reliability*, pages PMC2000–174, 2000.